DOCUMENT RESUME

ED 242 922 CE 038 715

AUTHOR Burnside, Billy L.; And Others
TITLE Training Feedback Handbook. Research Product 83-7.
INSTITUTION Army Research Inst. for the Behavioral and Social Sciences, Alexandria, Va.
PUB DATE Jan 83
NOTE 88p.; Prepared at the ARI Field Unit at Fort Knox, Kentucky.
PUB TYPE Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; Data Analysis; Databases; *Data Collection; Decision Making; *Educational Assessment; Educational Quality; Evaluation Criteria; Evaluation Methods; *Feedback; Followup Studies; Guidelines; Information Dissemination; Interviews; *Military Personnel; *Military Training; Postsecondary Education; Questionnaires; Records (Forms); Research Design; *Research Methodology; Surveys; Trainers

ABSTRACT

This handbook is designed to assist training developers and evaluators in structuring their collection of feedback data. Addressed first are various methods for collecting feedback data, including informal feedback, existing unit performance records, questionnaires, structured interviews, systematic observation, and testing. The next chapter, a discussion of integration of methods, deals with criteria for method selection, sampling techniques, and the structure of feedback collection. Covered next are the following aspects of managing feedback: automation of data management, organization of data, integration of data, data analysis and decision making, reporting of data, and follow-up. Appendixes to the guide include sample questionnaires, interview forms, observation forms, and a list of references. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED242922

CE

Research Product 83-7

TRAINING FEEDBACK HANDBOOK

ARI FIELD UNIT AT FORT KNOX, KENTUCKY

January 1983

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

AD-A131 715

U.S. ARMY RESEARCH INSTITUTE for the BEHAVIORAL and SOCIAL SCIENCES

Approved for public release; distribution unlimited.

U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Technical Review by:

Otto Kahn
Gary Sarli

NOTICES

FINAL DISPOSITION: This Research Product may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: This Research Product is not to be construed as an official Department of the Army document in its present form.

BEST COPY AVAILABLE

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Research Product 83-7	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Training Feedback Handbook	5. TYPE OF REPORT & PERIOD COVERED: Final Report September 1982-January 1983
6. PERFORMING ORG. REPORT NUMBER: --	
7. AUTHOR(s): Billy L. Burnside, Bob G. Witmer, & Donald M. Kristiansen	8. CONTRACT OR GRANT NUMBER(s): --
9. PERFORMING ORGANIZATION NAME AND ADDRESS: US Army Research Institute - Ft Knox Field Unit, Steele Hall, Fort Knox, KY 40121	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263743A794
11. CONTROLLING OFFICE NAME AND ADDRESS: US Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	12. REPORT DATE: January 1983
	13. NUMBER OF PAGES: 87
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): --	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
--

18. SUPPLEMENTARY NOTES
--

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Training Feedback	Training Management
Training Evaluation	Training Development
Training Effectiveness	Data Collection Methodology

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This handbook is designed to assist training developers and evaluators in structuring their collection of feedback. Six methods of collecting feedback are described, and practical guidelines for their application are offered. Issues in the management and analysis of feedback are also discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Research Product 83-7

TRAINING FEEDBACK HANDBOOK

Billy L. Burnside
Bob G. Witmer
Donald M. Kristiansen

Submitted by:
Donald F. Haggard, Chief
ARI FIELD UNIT AT FORT KNOX, KENTUCKY

Approved as technically adequate
and submitted for publication by:
Harold F. O'Neil, Jr., Director
TRAINING RESEARCH LABORATORY

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

January 1983

Army Project Number
2Q263743A794

Education and Training

Approved for public release; distribution unlimited.

FOREWORD

The ARI Fort Knox Field Unit has for several years been involved in the application of behavioral science to increasing the quality of the products of Army Centers/Schools. These products include training programs, training support materials, and, ultimately, trained soldiers who have participated in the programs or utilized the materials. A crucial aspect of quality control of these products is the gathering of valid and reliable feedback to support product evaluation and redesign.

For approximately three years, the Training Evaluation/Feedback Team of the Fort Knox Field Unit has addressed the collection and use of feedback by Center/School personnel. This effort has focused upon the design of comprehensive but practical systems for gathering feedback, and it has included several applications of the methods developed. The experiences gained during this applied research serve as the basis for this handbook.

The present report is designed to serve as a practical guide for the collection and management of feedback. It presents flexible guidelines which can be tailored to meet the needs of each feedback situation. It is intended primarily to serve the needs of training developers and evaluators within TRADOC Centers/Schools, but it should be of use to all who are involved in the evaluation of training.

EDGAR M. JOHNSON
Technical Director

TRAINING FEEDBACK HANDBOOK

BRIEF

Requirement:

Training developers and evaluators need feedback in order to determine whether products of the training development process meet the needs of users. This handbook is designed to assist these personnel in structuring their collection and use of feedback.

Procedure:

The numerous methods for collecting feedback are grouped into six categories in this report: informal feedback, existing unit performance records, questionnaires, structured interviews, systematic observation, and testing. Based upon available literature and previous research experience, the positive and negative aspects of each of these methods are discussed and general guidelines are offered for the application of each method. Other issues addressed include the integration of methods and the management and analysis of feedback.

Findings:

No one method is sufficient in and of itself for collection of feedback in all situations. But use of a mix of methods and tailored application of the guidelines offered in this handbook should provide the feedback that training developers and evaluators need. Integration of data to insure their accuracy, automation of data management, and follow-up to insure that data have the appropriate impact are key issues in the use of feedback.

Utilization of Findings:

This handbook will be useful to all training developers and evaluators in their collection of feedback from personnel in field units and in Centers/Schools. It will also have general utility for all personnel involved in the evaluation of training.

CONTENTS

	Page
SECTION I - INTRODUCTION	1
SECTION II - FEEDBACK METHODS	6
Informal Feedback	6
Existing Unit Performance Records	9
Questionnaires	14
Structured Interviews	20
Systematic Observation	24
Testing	30
SECTION III - INTEGRATION OF METHODS	34
Criteria for Method Selection	34
Sampling	36
Structure of Feedback Collection	38
SECTION IV - MANAGING FEEDBACK	39
Automation of Data Management	39
Organization of Data	40
Integration of Data	41
Data Analysis and Decision-Making	42
Reporting of Data	44
Follow-Up	45
APPENDIX A - SAMPLE QUESTIONNAIRES	47
APPENDIX B - SAMPLE INTERVIEW FORMS	61
APPENDIX C - SAMPLE OBSERVATION FORMS	71
APPENDIX D - REFERENCES	81

SECTION I - INTRODUCTION

In the context of Army training, feedback is any information about the results of the training development and implementation process. This information may take various forms, such as the results of tests administered to personnel enrolled in a training program, observations of training exercises, or opinions of participants in the training process. But regardless of its form, the purpose of feedback is the same - to provide input to evaluation of the effectiveness and efficiency of training. Analysis of feedback indicates whether or not the training process has achieved its goals. If the goals have not been met, then feedback should indicate the changes that are needed in training. Feedback is thus crucial to the training process; without it, there is no way of determining when training is effective.

Feedback is needed by all participants in the training process. Trainers and trainees need feedback to determine when training has been effective and when further training is needed. Training managers and commanders need feedback to determine the training readiness of their units and to schedule and provide resources for needed training. Training developers and evaluators in TRADOC Centers/Schools need feedback from users to determine whether the school's products have resulted in effective training, and to make modifications as needed. It is not possible to directly address the feedback needs of all these personnel in this handbook. Therefore, the emphasis here is upon the needs of one specific group, training developers and evaluators in TRADOC Centers/Schools. But the guidance and methods presented should be of general use to all individuals involved in the training process, including those managing training in field units.

This handbook is designed to assist training developers and evaluators in structuring their collection and use of feedback. Training development and evaluation are intertwined processes, but in most TRADOC Centers/Schools the personnel performing them are divided into two separate directorates. This is done to insure that personnel are not placed in the potentially compromising position of having to evaluate their own products. Training developers are personnel who produce programs or materials that are utilized to directly or indirectly support the training of soldiers. Here the term refers generally to personnel assigned to the Directorates of Training Developments (DTD's) of TRADOC Centers/Schools, but it may also include personnel in instructional departments and other agencies who are involved in the training development process. The term training evaluators is used here to refer generally to personnel assigned to the Directorates of Evaluation and Standardization (DOES's) of Centers/Schools, although it may refer to anyone who determines the effectiveness of training development and implementation. In the typical Center/School, training evaluators have the primary responsibility for collecting feedback and providing it to training developers and others having a need for it.

As stated earlier, feedback relates to the results of training development and implementation. From the perspective of training developers and evaluators, the results of training development are training programs and training support materials, and the results of implementation are the states of individual and collective training achieved through use of these products. Training programs are structured sets of procedures whose primary purpose is to provide needed

skills and knowledge to soldiers. Examples are basic entry training, basic and advanced NCO and officer courses, and training exercises derived from the Army Training and Evaluation Program (ARTEP). Training support materials are products of the training development process which enable training programs to be conducted. Included here are written manuals and guides, audiovisual materials, results of front-end analyses, ARTEP documents, and all other materials whose primary use is to support the training of soldiers.

In order to fully evaluate the effectiveness of training programs and training support materials, feedback is needed from all the users of these products. These users include commanders, training managers, trainers, and trainees, both in the institution and in operational field units. Feedback from personnel within the institution is designated internal, and feedback from personnel in field units is designated external. These two classifications of feedback are discussed jointly in this handbook, since the same general methods and guidelines apply to both.

Feedback from users should not be based largely on their opinions about products of the training development process, but should instead represent objective results of their experiences with these products. The ultimate purpose of training development products is to enable training to be provided to soldiers, and thereby to increase the combat readiness of the Army. Feedback should thus communicate the levels of training readiness resulting from the implementation of these products. It should indicate which products or parts of products have resulted in effective training, and which have not. In cases where effective training has not occurred, feedback should be detailed enough to allow determination of whether the problem lies in the product itself or in the way in which it was implemented. Corrective action can then be directed at the appropriate point in the training development and implementation process.

While feedback based upon the results of implementation of products is of prime importance, feedback is needed during all phases of the training development process. Training in the Army is based upon the Instructional Systems Development (ISD) model, which consists of five phases: evaluation, analysis, design, development, and implementation. Evaluation is not a separate independent phase in this model, but is rather a process which should be integrated into all the other phases. Feedback, or the data utilized in evaluation, can thus be useful at any point in the training development process. The decision to initiate the analysis and design phases may be based upon feedback indicating that new training programs or support materials are needed. Internal feedback gathered during the design and development phases may indicate that the appropriate tasks were not selected for training during the analysis phase, and the analysis may thus have to be revised. Feedback should flow continually throughout the training development process and not be gathered at just one point in time. In this handbook emphasis is placed upon feedback gathered during and after the implementation of products, but the methods and guidelines can and should be applied during all phases of the training development process.

The first step in the feedback process is to determine what information is needed. While the specific needs will vary from situation to situation, the

general characteristics needed in feedback can be summarized here. Training developers need feedback which is specific to each of their products. They need data indicating the proficiencies of soldiers and units who have participated in training programs or used training support materials, and information on how the programs were conducted or how the materials were used. They need to know which materials or parts of materials have been used in field units, and which had to be modified before they could be used. They need to know whether the tasks, conditions, standards, and resource requirements included in training programs and materials have proved to be realistic. Training developers and evaluators basically need all the feedback they can get, as long as it is manageable and related to specific products. Training developers within Centers/Schools are sometimes isolated from the realities of the field unit environment. They need to know what is happening in the field, and the collection of feedback and contacts with field personnel are the best ways to meet this need.

As indicated above, training developers and evaluators need many different types of feedback. They should obtain a variety of complementary data from different sources, in order to insure the accuracy of feedback. There are also numerous characteristics which each individual type of feedback should have. Feedback should be as specific and detailed as possible; it should relate to the performance of specific tasks or the use of specific parts of training materials. This is necessary so that the identification of problems in feedback leads to possible solutions. General feedback such as "lieutenants can't read maps" or "mechanics can't troubleshoot" is not of much use to training developers. In order to modify training programs and materials, the developers need more specific information detailing which personnel are having difficulty with which aspects of which tasks. Feedback should also be objective or based on observable evidence, rather than subjective or based just on opinions. Guidelines such as those offered in this handbook should be followed to maximize the accuracy of all feedback collected. Feedback should also be timely and representative of the current situation in the field. In today's rapidly changing state of technology, the situation in the field can change quickly and feedback processes should be attuned to this. Depending on the situation, feedback may need to be collected on each of several occasions, in order to insure that it is up-to-date. Since a high volume of feedback is needed, it should be quantifiable or easily encoded so that it can be entered into automated data processing systems. No feedback should be collected without firm plans as to how it will be analyzed, managed, and used. Use of this handbook will not guarantee fulfillment of all the feedback needs discussed above. But sound refinement and tailoring of the principles offered in it should help structure feedback to meet the needs of the Army.

One characteristic of feedback which deserves special mention here is the need for it to be coordinated and integrated. One aspect of this is the need for training developers and evaluators to communicate continually. As was pointed out earlier, training developers and evaluators at most TRADOC Centers/Schools are divided into two directorates, the DTD and the DOES. Some independence is needed between these two sets of personnel in order to lessen the probability of bias during the evaluation of products, but they should

constantly communicate with each other as to what feedback has been collected and what further feedback is needed. The feedback process should be a continuous loop, with training developers identifying what feedback is needed, training evaluators collecting it and providing it to those who need it, and training developers indicating what was done in response to the feedback and what new feedback is needed. Training evaluators gather feedback during Branch Training Team (BTT) visits or other visits to field units. They should involve training developers in planning these visits and specifying the feedback to be collected. A general telephonic or written request to training developers for questions to be asked during upcoming visits may not be sufficient. Face-to-face active coordination between training developers and evaluators is needed prior to feedback collection visits. Consideration should be given to continuing this coordination throughout field visits by including training developers on data collection teams. In this way training developers may be better able to collect the detailed feedback that they need and may have more confidence in the data collected. Coordination should also continue during analysis and use of feedback, in order to lessen the probability of differing interpretations of the data collected and to insure a common understanding of the feedback needed during future field visits. These continuous coordination efforts should eliminate duplications of effort by training developers and evaluators and should insure that feedback meets the needs of all concerned.

Another aspect of the coordination of feedback involves the integration of the methods used to collect it. As will be discussed in detail in the next section, there are several methods through which feedback can be collected, and none of them are sufficient for all situations. A mix of methods must thus be tailored to the resources available and the scope of the problems encountered. For example, feedback on a problem with a training product may first be provided to Center/School personnel through informal communications from the field. The problem can then be further examined through administration of surveys or interviews to field personnel. If the problem appears to be a severe one, it may then be necessary to address it further through structured observations or testing in the field. The integration of different methods may also increase the accuracy of feedback, since the data collected from different approaches or sources can be used as cross-checks. Using the appropriate mix of methods in collecting feedback is an important concern, and a separate section will be devoted to it in this handbook.

The final aspect of the coordination of feedback to be introduced here involves the integration of the data collected. All the data relating to a particular product or issue should be managed so that they can be used in a complementary fashion. Feedback should be maintained in a centralized data base which is accessible to both training developers and evaluators. In this way duplications of effort can be reduced and decisions to revise training products can be based upon all the data available. Guidelines for the management and use of feedback will be presented in later sections of this handbook.

The collection and use of feedback is a continuous adaptive process, and there is no unique solution or set procedure for accomplishing it. Thus, while the goal of this handbook is to increase the structure and standardization of the

flow of feedback to and within TRADOC Centers/Schools, it presents general guidelines rather than rigid procedures to be followed. General "how to" procedures are presented which training developers and evaluators can tailor to their individual organizational needs and available resources. Included in these procedures are descriptions of the primary methods which can be used to collect feedback, discussions of general approaches to analyzing and managing the data collected, and suggestions for ways to utilize feedback to improve the products evaluated.

SECTION II - FEEDBACK METHODS

There are many ways in which training developers and evaluators can collect feedback relating to products of the training development process. External feedback can be gathered through formal visits to field units by Branch Training Teams (BTTs) or other groups of Center/School personnel, through communications with field personnel using telephones or the mail, or through interaction with field personnel when they are temporarily available at the Center/School. Internal feedback can be gathered through visiting institutional training sites or through interactions with trainers and trainees. Several general administrative issues arise during planning for the collection of feedback, such as who should collect it and how often it should be collected. Such general issues are not directly addressed in this handbook, except in terms of broad guidelines. Rather, this chapter of the handbook focuses upon relatively specific feedback collection methods which can be applied either at the institution or in the field. Managers of BTT's or other feedback collection efforts should integrate these methods and tailor the guidelines offered to meet the needs of each particular situation. In this way, structure can be introduced into the feedback process while maintaining its needed flexibility.

The feedback methods discussed in this chapter are organized into six categories: informal feedback, existing unit performance records, questionnaires, structured interviews, systematic observations, and testing. Each of these methods is described below, followed by a discussion of its positive and negative aspects and guidelines for when and how it should be used. In the cases of questionnaires, interviews, and observations, sample forms for implementation of each method are included in an appendix. The forms are intended to be generic examples of the application of guidelines presented, and they are not intended to be extracted and used, as is. The forms used in any feedback collection effort will likely have to be tailored to the needs of the particular situation. Not all the guidelines and information relevant to each feedback method can be presented in this handbook. The major points are covered here, and relevant references for several of the methods are listed in Appendix D. Training developers and evaluators planning to make extensive use of a particular method should read the relevant references, in addition to this handbook. Since no method is totally sufficient in and of itself, the integration of methods into a total feedback system is discussed in the section following this one.

Informal Feedback

Informal feedback is that which is unsolicited and flows through unstructured means. This includes information that is transmitted through word of mouth and through general written formats, such as letters and completed DA Forms 2028 recommending changes in training support materials. Numerous examples of feedback fall under this category. To provide external feedback, training managers in the field may telephonically or in writing contact personnel at the appropriate Center/School about a problem with training support materials or the training of soldiers recently arriving in field units. In

some cases, special "hot lines" to the Center/School are set up for this purpose. Or field unit personnel may contact training developers about perceived problems while they are visiting the Center/School to attend a conference or for some other purpose. Almost all training support materials include a request for comments to be sent to the developer, either on a DA Form 2028 or on a questionnaire form provided with the publication. Requests for general information to be sent using DA Forms 2028 are included here as examples of informal feedback. Structured forms provided with publications for obtaining feedback are addressed in the later discussion of questionnaires.

Internal feedback may also be transmitted informally. For example, a trainer in an instructional department may informally contact personnel in the DTD with comments on the Program of Instruction (POI) developed for a particular course. Or subject matter experts within the Center/School may informally review and comment upon training programs falling within their area of expertise. The turnover of military personnel within DTD's provides another means of obtaining informal feedback. These individuals may recently have been transferred from a field unit assignment, and thus be in a position to comment upon the usefulness or likelihood of acceptance of particular materials in the field.

Another characteristic of informal feedback which should be mentioned here is that it is usually based upon unstructured personal observations. These may be the observations of training managers in the field during the conduct of training, or the observations of Center/School personnel visiting field units and observing training without applying structured feedback collection methodologies. For example, general officers from the Center/School may observe training while visiting field units and provide comments in memoranda to institutional agencies or letters to the field. Internal feedback may be gathered through informal observation of the initial administration of a new training program, or through gathering observations from institutional trainers. Informal feedback is thus a by-product of the direct, personal experiences of personnel with the products of the training development process.

Positive and Negative Aspects

The primary positive aspects of informal feedback are that it requires relatively few resources to collect, it does not detract from or interfere with training, it is usually based upon direct experiences with the products evaluated, and its open-ended or unstructured nature may lead to new and unexpected information. Informal feedback may uncover unanticipated problems with products that training developers and evaluators would never have thought to address, otherwise. These advantages are significant enough to support the continued use of informal feedback, but several potential disadvantages should be kept in mind.

One negative aspect of this type of feedback is that it is not as readily available as one might expect. Probably due to time constraints, training managers in the field do not frequently provide unsolicited feedback to training developers and evaluators. The flow of DA Forms 2028 into DTDs could best be

described as a trickle. A feedback system which depends heavily upon informal information is thus likely to suffer from a lack of data. The provision of informal feedback depends largely upon the attitudes of trainers, training managers, and commanders and the number of informal contacts they have established at the Centers/Schools. Some field personnel may have many acquaintances at a Center/School and feel free to contact them regularly, while others may be hesitant to surface problems outside their units. The informal feedback available may thus over-represent the observations of particular units and individuals and not provide an accurate picture of the overall situation in the field (i.e., the squeaky wheel gets the grease). No structured sampling techniques are applied in gathering informal feedback, so there is no guarantee that the samples obtained are representative ones.

Another problem with informal feedback is that it tends to be general in nature and does not identify specific tasks or parts of training support materials where problems lie. It is thus usually necessary to follow up this type of feedback with methods which yield more specific information. There is also usually no guarantee that informal feedback is accurate. It is largely based upon personal observations, and these may be over-generalized and influenced by interpretations and opinions. This type of feedback should thus not be accepted as being definitive, but rather as initial indications of potential problem areas that need to be addressed further. A final problem with informal feedback is that its lack of structure makes the data obtained difficult to tabulate, analyze, and integrate. The maintenance of the data may thus be cumbersome, since they cannot easily be fit into automated data files.

Guidelines for Use

In spite of the problems discussed above, informal feedback should be a part of the evaluation process. Informal communications should be encouraged among all participants in the training development process, since useful information may be transmitted in this way. But informal feedback should not be relied upon to provide definitive information to support final decisions about revisions needed in training development products. Guidelines for use of informal feedback in the overall evaluation process are presented below.

1. Since informal feedback tends to be general, personnel receiving it should gather all the details they can at the time of input. If a problem with a product is noted, then all the details of the situation in which that problem occurred should be identified. If a problem with training support materials is noted, then the specific parts of the materials needing revision should be identified. The number of personnel noting the problem should also be identified, along with general demographic data indicating their background and experience level. These details should be recorded in a written format and filed in an accessible manner. Trusting informal feedback to memory may lead to distortions in the information obtained.

2. While informal feedback is by definition unsolicited, it can and should be solicited in general ways. The setting up of "hot lines" in Centers/ Schools for the provision of informal feedback should be encouraged, and the telephone

numbers of training developers and evaluators should be made widely available to field unit personnel. Command letters and other communications distributed by Centers/Schools should encourage the return of feedback. Personnel attending conferences or short courses at a Center/School should be encouraged to contact training developers and evaluators with feedback on the use of training products.

3. When informal feedback is obtained, it should be responded to, whether or not it led to a revision in a training product. Those providing feedback should be told what was done about problems they identified or recommendations they offered. If nothing was done, they should be told why. In this way, personnel can become confident that feedback they provide will be attended to and that appropriate actions will be taken.

4. Informal feedback should be used only as an initial indication of where problems may lie. It should always be followed up by collection of further feedback using methods discussed below, in order to determine its scope and accuracy. Revisions in training products should not be made based upon feedback that may be the opinion of one person. Informal feedback can be an important starting point in the feedback process, but it is not the end point.

Existing Unit Performance Records

Unit personnel maintain numerous records of the results of the implementation of training programs in their units. These records are used to manage training (e.g., to determine in what areas further training is needed) and to determine the combat readiness of individuals and units. A few of the numerous examples of such records are described below, to serve as a basis for discussion of their usefulness as feedback to Center/School personnel.

In field units, written reports of the results of the unit's performance on recent formal ARTEP exercises are usually maintained at battalion or squadron level. The results usually consist of an overall rating, narrative descriptions of general performance strengths and weaknesses, and criterion-referenced ratings of the performance of each company-sized unit on various missions. Records of individual task performance are generally maintained in the form of results of recent Skill Qualification Tests (SQT's). Battalion-level summaries of these results indicate the percentages of soldiers who correctly answered each question on the written portion of the test and who properly performed each task on the hands-on portion. The results of various qualification exercises are generally maintained at battalion or squadron level, and they span the individual/collective task dimension. These results may provide indicators of training effectiveness at the level of individuals (e.g., firing of the M16 rifle), crew (e.g., firing of tank weapons on Table VIII), and platoons (e.g., control of firing and movement of tanks on Table IX). The indicators usually consist of numbers or percentages of individuals or units qualified, and perhaps general reasons for non-qualification. Other records indicating the effectiveness of training may be maintained at various levels in field units. For example, job books maintained by supervisors are designed to provide a record of individual performance on specific tasks, along with narrative comments relating

to training needs or problems. Such records are usually not consolidated at battalion level.

Records are also maintained on institutional training which provide measures of the effectiveness of implemented training programs. Results of written and hands-on performance tests given during or at the end of a course of instruction indicate the proficiencies of soldiers on specific tasks. Records of the numbers of soldiers failing to complete particular training programs and the reasons why provide more general feedback on the effectiveness of training. Results of internal inspections conducted by training evaluators may indicate whether lesson plans and other training support materials are being used as they were designed to be and whether other general problems exist in the training base.

Most of the sorts of records described above are designed to meet the feedback needs of training managers within units. They provide these personnel with important information which can be used to distribute training resources where they are needed. The key question to be addressed below is whether these records can also meet the feedback needs of training developers.

Positive and Negative Aspects

A potential advantage in the use of existing unit records as feedback to training developers is based upon the assumption that these records are readily available. Few resources should be required to gather data that are already available in units, and the provision of such feedback should place little burden on unit personnel, since they would not be required to generate any data in addition to that normally produced. The problem is that this assumption is often not valid; unit performance records are often not readily available in units. For example, ARTEP results may not be available because the unit has not participated in ARTEP exercises during the past year, or has participated in nonstandard ARTEP exercises. SQT results may not be available for all MOS's, expecially since conduct of a hands-on SQT is now optional. Job book entries may be incomplete, since supervisors do not have time to record daily observations of task performance by all soldiers working for them.

The unit records that are available may not be in a standard format that is comparable from unit to unit. For example, ARTEP's and qualification exercises are conducted differently in different battalions. Range and ammunition availability affect the way that major exercises are conducted, such as Table VIII in tank battalions. Results of such exercises cannot be compared and integrated across battalions without consideration of the conditions under which the exercises were conducted. Different scenarios and conditions are employed in ARTEP exercises in different battalions, making it difficult to combine results of such events across units to determine the effectiveness of training programs in the field as a whole. Training developers and evaluators should not assume that unit performance records are readily available in a form that can be integrated across units. The acquisition and analysis of unit performance records is a more difficult task than might initially be expected.

Another possible advantage to using existing records as feedback to Centers/Schools lies in the assumption that these records represent objective and reliable measures of unit performance. This assumption is also not always valid. Resource and time limitations in the field often do not allow completely objective evaluations of all aspects of unit performance. Unit performance records are thus based to some extent upon the subjective judgments of commanders and training managers serving as evaluators. For example, ARTEP's and other collective exercises are complex, integrated events, and evaluation of performance on them is to some extent subjective. Control of fire and movement techniques are examples of collective tasks which are usually evaluated subjectively. Unit performance records reflect direct experiences with the implementation of training programs, but real-world limitations frequently prevent these records from reflecting such experiences in the objective detail needed by training developers. Such detail can be increased by providing evaluators with extensive training in evaluation and the feedback of information, and by using performance records based upon data gathered using relatively objective methods, such as the Multiple Integrated Laser Engagement System (MILES) or other instrumentation available at the National Training Center (NTC). But resource constraints often limit the training that can be given to evaluators and the availability of MILES or NTC, so the utility of existing performance records on feedback to training developers is frequently limited.

A major disadvantage with the use of unit performance records as feedback to Center/School personnel is the fact that these records are often general rather than task-specific in nature. For example, results of ARTEP exercises usually delineate major strengths and weaknesses of participating units, but they do not point out specific problems with training programs or specific tasks for which further training is needed. Results of qualification exercises indicate the percentage or number of individuals or units qualified, but may not delineate specific problems that occurred. General records of the results of exercises may be sufficient to meet the evaluation needs of unit personnel, since these records are supplemented by first-hand experiences gained during exercise conduct. That is, commanders and training managers may be able to identify specific areas or tasks for which further training is needed, based upon their observations and personal interaction with exercise evaluators. But these details are often not recorded in written form, and thus are not readily available as feedback to training developers. For example, during the conduct of an ARTEP exercise problems may be noted with the ARTEP document itself, but such observations are usually not recorded in written reports of ARTEP results. General performance records must thus be supplemented with first-hand observations or with interviews of experienced personnel. When unit performance results are recorded in task-specific detail, they are often highly perishable and are not available for a long period of time. Memories of individuals participating in exercises are also highly perishable and become general and less accurate over time. The gathering of feedback based upon the performance of unit exercises should thus be coordinated with unit training schedules so that it can be accomplished during or soon after exercise completion.

There are some individual performance records available which are sufficiently detailed to allow the identification of specific changes needed in

training programs or in training support materials. SQT results have in the past provided measures of soldiers' performance on specific tasks. Training developers could use such results to identify specific tasks for which revisions in training are needed. However, the requirement for a regular, formal administration of the SQT has recently been eliminated, so these results are expected to be less available in the future. Results of tests administered during and at the completion of institutional training programs generally provide task-specific measures of soldiers' performance. Such results should be readily available to training developers for use in redesign of training programs. Measures of performance are generally more formal and more specific in institutional than in unit training programs, so existing performance records are likely to be more useful as internal rather than as external feedback. Within the institution, training developers should be able to check the accuracy of performance records and obtain needed further detail by visiting training and testing sites.

A final concern with the use of existing performance records as feedback to training developers is related to the purposes for which evaluation is conducted. Training developers and evaluators in Centers/Schools gather feedback for somewhat different purposes than do commanders and training managers in units. Center/School personnel need feedback to determine the effectiveness of training programs and training support materials, and to identify any revisions needed. They thus have no need to identify specific units or individuals from whom feedback is collected. Field unit personnel need feedback to determine the readiness of specific individuals and units, and to identify further training needed. Existing unit performance records are designed primarily to meet the evaluation needs of field unit personnel. These records are thus largely informal, highly perishable, and supplemented by first-hand experiences rather than by written detail. They often do not meet the evaluation needs of training developers, and redesign of them to meet these needs might reduce their usefulness to unit personnel. For example, knowledge that their performance records were to be provided as feedback to a Center/School might influence unit personnel to "scrub" such records to make certain that they would not reflect negatively on the unit. The objectivity of existing unit performance records can be questioned, and the provision of such records to higher headquarters might produce further problems here. Also, unit personnel do not have time to record and maintain performance data in the detail that they are needed by training developers. The management of such data might reduce the amount of time that senior unit personnel could spend in first-hand monitoring of training activities. Training could become a set of formal, discrete activities to meet the needs of Centers/Schools, rather than a continuous adaptive process to meet the needs of units. It is thus best that unit performance records remain oriented toward meeting the evaluation needs of unit personnel, and not be tailored to meet the needs of Center/School personnel. These records can provide some useful information to training developers, but they cannot in and of themselves meet all of these personnel's feedback needs.

As was suggested earlier, records based upon the conduct of exercises using MILES or instrumentation available at NTC may provide feedback which alleviates some of the problems discussed above. Such exercises generally

provide results which are relatively objective and detailed, so these data should be gathered as feedback whenever possible. However, limits on the availability of MILES equipment and restrictions on the use of NTC data may limit the extent to which results of tactical engagement simulation can be used as feedback. The use of simulations and automated data collection devices provides a promising approach to the efficient gathering of objective feedback in the future. But until such simulations and devices are routinely available, caution should continue to be applied in the use of existing records as feedback.

Guidelines for Use

As discussed above, there are many problems with the use of existing unit performance records as feedback to Center/School personnel. These records are often not readily available, not available in a standard form, not totally objective, and not available in sufficient detail. But since they do reflect the results of the implementation of training programs in at least a general sense, unit performance records can have some utility in a total feedback system. These records should be used as feedback in accordance with the guidelines presented below.

1. Since existing unit performance records are often not available in sufficient detail to meet the needs of training developers and evaluators, as much detail as can be used should be gathered to supplement these records. If the results of a qualification or ARTEP exercise are collected, details on how the exercise was conducted and scored should be gathered. Training developers and evaluators can accomplish this by interviewing exercise participants or by observing the conduct of the exercise. Since detailed performance records and individuals' memories are highly perishable, supplemental information such as scenarios used, problems with scoring or targets, and specific tasks or subtasks resulting in performance difficulties should be gathered as quickly as possible following exercise completion. Performance results should be gathered in as original a form as possible; e.g., original scoresheets should be gathered in addition to summary results. When observing the conduct of an exercise, training developers and evaluators should spotcheck the accuracy of scoring decisions and record factors which may influence the results obtained. Additional details should be obtained in as direct and timely a manner as is possible, and plans should be developed for use of these data prior to their collection.

2. When gathering unit performance records as feedback, the anonymity of the units and individuals involved should be insured. Identifiers of individuals and units should be removed from scoresheets and from all other data collection forms. Assurances of anonymity should be given to personnel in units from which performance results are collected. In this way, tendencies to bias results or cover up problems can be reduced. The feedback gathered should not be used in any way which might reflect negatively upon particular individuals or units.

3. Performance results are generally gathered more formally and in more detail in the institutional setting than in field units. The gathering of supplemental details is also less costly in the institution than in the field. Existing performance records should thus be more heavily relied upon as internal rather than as external feedback. Training developers and evaluators should obtain results of performance tests and exercises from training programs conducted within their Center/School. They should also visit institutional training and testing sites in order to gather further details and to check the accuracy of the data obtained.

4. As is the case with informal feedback, existing performance records should be used as initial indicators of where problems may lie in the training process, and not as final, definitive feedback. These records should be supplemented by interviews, observations, and other feedback methods, in order to insure the accuracy of the feedback obtained. Existing records should not be heavily relied upon in a feedback system, but should be used as an indication of areas in which further information is needed.

Questionnaires

Perhaps the quickest way to obtain information for feedback purposes from a group of respondents is to require them to complete a questionnaire. A questionnaire is any ordered set of questions or items designed to elicit written responses from a specified population or group. Broadly speaking, questionnaires may take on a variety of forms ranging from rating scales to multiple choice items to open-ended questions. Open-ended questions are ones that allow the respondent to answer in his/her own words, including whatever qualifiers he/she likes. Some forms lend themselves better to collecting accurate feedback than others, and these will be identified later in this section. There are a number of other factors influencing the accuracy of information obtained through questionnaires, which will also be discussed later in this section.

Although questionnaires can be used for gathering feedback in a variety of situations, the power of the questionnaire is best demonstrated when you have a large group to survey and few data collectors, or little time to gather the data. In such situations, the questionnaire may be the preferred method of obtaining feedback, and possibly the only method. Questionnaires can be administered to groups or individuals in the field or at the institution for obtaining quick feedback on personnel or training matters. Questionnaires are flexible in that they can be administered in person or through the mail. However, mailing out questionnaires is not recommended for reasons to be discussed later.

Positive and Negative Aspects

As mentioned above the questionnaire can be used to gather a considerable amount of feedback in a relatively short time with only a small investment in resources. A single data collector armed with multiple copies of a questionnaire can collect information from several hundred respondents in a relatively short period of time. If the questionnaire has been carefully designed, the

same data collector could, with the aid of computerized data management, analyze the data and provide feedback on the results in a matter of days.

In general, research has shown questionnaires to be a reliable way of gathering information. Unless the questions are worded ambigously or beyond the reading level of the respondent, questionnaire responses tend to be consistent with the intent of the questions. The consistency in answering questionnaires may in part be explained by the fact that each respondent reads the same questions, worded in the same way, in the same order on each occasion. This allows the collector of feedback to combine the responses of the individual respondents for arriving at general conclusions concerning the development and implementation of training products.

Questionnaires allow the training developer or evaluator to gather information that might be extremely difficult to see or might take months to obtain through direct observation. In essence, the questionnaire expands one's ability to observe by letting the respondent act as the eyes and ears of the would-be observer. While this expanded ability to observe makes the questionnaire extremely efficient, it does not guarantee that responses to the questions will be accurate. Accuracy may be adversely affected by the inability of the respondent to recall relevant information or by the respondent's unwillingness to answer questions completely and honestly. Honesty and memory issues will be discussed further in the guidelines for using questionnaires at the end of this section.

The biggest drawback to questionnaires is that bad questions are easy to write. Too often questionnaires are written for no reason other than to meet an evaluation requirement. Often those writing the questionnaire have little idea of what information is needed or what they will do with the information once they obtain it. The result is the proliferation of a large number of questionnaires containing very general questions that produce data that are not useful for improving training programs or products. Such Questions as "What is the present state of training in your battalion?" or "How well trained are soldiers when they arrive in your unit?" are unlikely to provide information about the extent and effectiveness of the training conducted in the battalion or about the tasks the soldiers could or could not perform upon arrival in the unit. To obtain the latter information, more specific questions are needed, perhaps supplemented by observation and testing.

Lack of specificity in questions reduces the amount of useful information that they produce. To make matters worse many questions call for subjective responses, consisting largely of the respondent's impressions, opinions, or subjective evaluations of training programs, materials, or soldier proficiency. Responses are typically obtained in terms of five or six-point rating scales, which have been used in a wide variety of applications, often inappropriately. Some of the more common uses have been for obtaining task criticality ratings, relative task difficulty ratings, and estimates of soldier proficiency by supervisory personnel. Rating scales have also been used to evaluate training effectiveness, training materials, and in numerous other applications where there are

more accurate ways of obtaining the information required. Because rating scales are relatively easy to design and quick to administer and score, they tend to be overused. Even when used appropriately, the accuracy of information obtained via rating scales is adversely affected by rater errors. Among the most common rater errors are rater leniency, central tendency errors, and the halo effect. Leniency errors occur when raters avoid using the low extremes of a rating scale. These errors often arise because raters do not wish to give ratings that can be interpreted negatively. Other raters may be biased against giving extreme ratings, either high or low. When both high and low extremes are avoided, ratings tend to cluster around the mid-point of the scale. When ratings cluster around the mid-point, the differences among things rated may be so small as to be of little practical significance. The halo effect occurs when the rater bases all ratings on a global or overall judgment of the person or thing being rated. For example, a supervisor asked to rate a soldier's performance on specific tasks may rate him or her high on all tasks based on the global judgment that the soldier is a good worker. The supervisor may give equally high ratings on all tasks even though some of the tasks were never observed and the soldier does not perform all tasks equally well. This rating tendency detracts from the ability to discriminate between different aspects of performance.

The accuracy and usefulness of information obtained by questionnaires largely depends on the questions asked and how they are asked. Often the wording of a question may indicate the expected or preferred response. Such questions are referred to as leading questions. Two examples of leading questions are given below.

1. Which tank do you prefer, the faster, lower profile M1 or the larger, slower M60?

2. Do you feel the weapon systems are not reliable, and, more frequently than not, do not operate to the full potential for which they were designed?

Such questions tend to support the preconceived notions of the questionnaire developer rather than to obtain accurate feedback.

Another problem with questionnaires is that they are unpopular with those persons from whom the feedback must be obtained. The unpopularity of questionnaires increases as the number and length of questionnaires to be completed increases, due to fatigue and other factors. Field personnel perceive, often correctly, that the questions asked have no real purpose and that no meaningful changes will be made as the result of their responses. Thus they do not devote as much thought to the questionnaires as they could.

Guidelines for Use

Because they provide an efficient means of obtaining information, questionnaires should be considered whenever there is a need for timely feedback from the field and resources are limited. Although no feedback effort should rely on questionnaires alone, well-designed and carefully administered questionnaires can provide a substantial amount of useful information. To help in the design

and administration of questionnaires, the following guidelines are offered. Following these guidelines should improve the usefulness and accuracy of the information obtained. Two sample questionnaires exemplifying many of these guidelines are included at Appendix A.

1. No questions should be written until the questionnaire developer identifies what information is needed and how it will be used. When it has been determined what information is needed and how it is to be used, the next step is to determine if questions can be formulated to obtain the desired information. Given a question that is likely to yield accurate information, the feedback collector must next identify the person or group who is in the best position to provide this information. Factors to be considered in selecting those who will provide the information include whether they might have direct knowledge or recent experience pertaining to the item and the likelihood that they will answer the question openly and honestly. To encourage openness and honesty, the person responding to the questionnaire and his/her unit should not be identified whenever possible. If the respondent's anonymity cannot be insured, then additional information supporting the accuracy of the questionnaire responses may have to be obtained. If the questionnaire asks for information of a sensitive nature and responses are not anonymous, then Privacy Act considerations come into play.

2. Questions must be worded so that they are direct and to the point. They should be readily understood by the respondents for whom they are designed. The questionnaire writer must be careful to use terms that mean the same to each respondent as they do to the writer. Ambiguous questions that depend heavily for an answer on the respondent's own frame of reference should be reworded or omitted. For example, the question "Are crews selected to fly together most of the time?" is ambiguous and may be answered yes or no depending on the respondent's interpretation of the phrase "most of the time." The question might better be worded in terms of how often or what percentage of the time crews fly together. Questions inquiring about the adequacy, sufficiency or satisfactoriness of this, that, or the other tend to be ambiguous, as do questions asking whether certain events occur frequently, occasionally, sometimes, regularly or often.

3. To be really useful feedback must be specific, and this requires the use of specific questions. Questions beginning with "how do you feel about", "what is your opinion of", "what do you think of", and similar phrases are unlikely to produce very much useful feedback. Responses to such questions will vary widely from one respondent to the next, making it difficult to draw any conclusions from the respondents as a group. Furthermore, questions phrased this way are more likely to elicit general opinions than to generate specific factual information. To obtain specific factual information, questions must precisely specify what information is required Instead of asking a tank gunner how he feels about the unit's gunnery training, you might ask him what livefire gunnery exercises he has participated in since arriving in the unit. Rather than ask the platoon leader what he thinks of a particular ARTEP document, you might ask him/her which parts of it were used during the most recent ARTEP, and then ask what problems were encountered in using those parts of the ARTEP document.

4. When open-ended questions are used, the questions should be sharply focused. Questions such as, "What equipment do you presently use in troubleshooting the M1 turret?" and "What equipment do you need for turret troubleshooting that you don't now have?" are examples of sharply focused questions. These questions are preferable to general open-ended questions, such as "What do you think of the equipment you have for troubleshooting the M1 turret?" Whenever possible, open-ended questions should be replaced by questions offering specific response alternatives. In the example above, the questionnaire developer may know in advance what equipment the mechanic has available for troubleshooting the M1 turret. Therefore, he/she can list the available equipment (e.g., STE M1, multimeter) as response alternatives and allow the respondent to check those that he/she uses in troubleshooting the turret. Providing response alternatives saves time in administering and scoring questionnaires and is highly recommended. Sometimes in pretesting a questionnaire on a small number of respondents the range of response alternatives to some open-ended questions will become apparent. When this occurs the response alternatives to these questions can be supplied in the final version of the questionnaire. Not every open-ended question, however, may be changed in this way since a given question may have too many possible response alternatives to list.

5. Certain types of questions should be avoided because they do not yield accurate, useful data. Other types of questions should only be used under very special circumstances and on a limited basis. Leading questions, in which the wording of the question implies an expected or preferred response (as in the examples given earlier), should be avoided. Also avoid phrasing questions in negative terms. Respondents often overlook the negatives in reading such questions and interpret the questions incorrectly. Negative questions are often leading questions as was illustrated in the example mentioned earlier. Do not ask two different questions as one compound question. For example, the question "Was the feedback you received complete and accurate?" actually consists of two separate questions (i.e., was the feedback complete? and was the feedback accurate?). Such compound questions should be avoided because a person can agree with one part of the question and disagree with the other. Rating scales and other subjective appraisal techniques should be used cautiously in questionnaires designed to provide feedback. Rating scales seldom provide the kinds of detailed information that are required in making specific changes to training programs and materials. If it is determined that a rating scale can provide the desired information, then appropriate guidelines offered in the references in the "Questionnaire" section of Appendix D should be followed. In addition, if rating scales or other subjective appraisal techniques are used, the accuracy of these techniques should be checked by comparing them with observational or performance data prior to using them to collect feedback. The only exception to this rule is when the data of interest are the appraisals, judgments, or perceptions themselves (e.g., judgments concerning task criticality or measures of job satisfaction). Generally subjective appraisal techniques should not be used for determining soldier proficiency, or effectiveness of training programs and materials unless it can be clearly demonstrated that the subjective measures accurately reflect proficiency and effectiveness as measured by more objective methods.

6. If rating scales are to be used, make the scales as explicit as possible. Phrase the scales in terms of explicit observable measures of performance, rather than in vague general terms such as "average", "below average", etc. Describe each rating point in terms of the behavior that it represents. Consider asking raters to provide specific examples to support their ratings. Be sure that the raters have the experience or background to rate what is being rated. Give raters the option of indicating that they have not had experience relevant to the items being rated. Train raters in making subjective appraisals. Rater training should include experience with the rating scales to be used, a discussion of common errors (leniency and halo effects) associated with ratings, and a discussion of the dimensions of the situation being evaluated.

7. Questionnaire items should be obvious and straightforward. There should be no need for the respondent to stop and figure out what the question means or to search his/her memory for long-forgotten facts. Questions should deal primarily with events that have occurred in the last few days or weeks, or with easily remembered events. In forming questions which depend on a respondent's memory or recall capabilities, the time period covered by the question must be carefully defined; the "when" should be specifically provided. Rather than ask the respondent what training he/she has received recently, ask how many hours of MOS training he/she received in the last three days or last two weeks. Do not expect to get accurate, detailed accounts of events that occurred many months or years ago. Questions can address events occurring months or even years ago when the event is likely to be especially significant or salient to the respondent, and absolute accuracy and fine details are not required.

8. Questionnaire items should be arranged in a logical sequence, both to facilitate the respondent's recall of the information and to shorten the time needed to answer the questions. In general, questionnaires should require no more than an hour of the respondent's time for completion. Shorter questionnaires are recommended. Do not pad the questionnaire with additional items just to make it longer. Do not include questions that elicit information that is only of mild interest. Questions included should be those that produce data having a known use. Unless a use for the data produced by a question can be specified in advance, the question should probably be discarded.

9. The questionnaire should have a title or heading to identify it on each page. If the questionnaire consists of more than two pages, the pages should be numbered. Each item included in the questionnaire should also be numbered. Be sure to leave enough room under each question for the respondent to write an answer to the question. While the order of the questions is not critical, a logical arrangement of items by subject matter or chronological order facilitates questionnaire administration. Questionnaires should usually include written directions for completing the questionnaire. Directions should always be clear and prominently displayed. Generally, respondents should be told the purpose of the questionnaire either verbally by the test administrator or as part of the written instructions.

10. Distributing and receiving questionnaires through the mail is not recommended. The return rate of mailed-out questionnaires is typically so low that

little confidence can be placed in conclusions drawn from the select few that are returned. Much less thought and effort is usually devoted to a mailed questionnaire than would be devoted to an interview. While this statement probably holds true of questionnaires in general, questionnaires received in the mail probably receive even less attention than other questionnaires. There is much to be said for administering questionnaires in person. Not only do you get more completed questionnaires, but the effort invested by respondents in completing the questionnaires is generally greater. Also the administrator is available for handling any questions that arise about the questionnaire items and for checking each questionnaire for completeness as it is returned. When a questionnaire is administered to a group, all respondents should be kept at the site where the questionnaire is administered for a predetermined period of time or until all respondents have completed the questionnaire. Respondents should not be permitted to leave the site early.

11. All newly developed questionnaires should be pretested prior to using them to gather feedback. The questionnaire should be pretested on the same type of respondents as those for whom the questionnaire was developed. For example, if the questionnaire was designed for tank commanders in an M1 battalion, then it should be pretested using M1 tank commanders. In pretesting the questionnaire, have the respondents read each question, explaining what it means in their own words. Then ask the respondents to answer the questions. By noting instances in which questions are not understood or misinterpreted and which questions produce little information, the questionnaire can be refined. Such pretesting may result in rewording some questions, focusing others more specifically and eliminating still others. Pretesting may also suggest additional questions to be included in the final version of the questionnaire.

Structured Interviews

Another efficient method of obtaining information for feedback purposes is through the use of structured interviews. A structured interview consists of a prepared set of questions verbally asked, on a one-to-one basis, of one or more interviewees by a skilled interviewer. Typically, the same questions are asked in the same order and in the same way each time the interview is conducted.

Structured interviews usually work best when the information gatherer has a clear idea of the kind of information he/she is seeking. For training developers looking for specific information about particular products, the method is therefore ideal. Interviews may be conducted at the institution with instructors or with soldiers just prior to their graduation from any given training program. Interviews may also be conducted with soldiers coming back from the field to attend a resident course at the institution. Soldiers returning to the institution for training are an especially good source of information for training developers who may otherwise have little contact with personnel in the field. During BTT visits, training developers and evaluators can interview unit personnel to assess the state of training in the unit and to evaluate the products of the training development process. The information thus obtained can be used both by unit leaders to better manage unit training and by training developers at the institution to improve their training products.

Positive and Negative Aspects

Although structured interviews require more time and effort than questionnaires, the payoff in terms of quantity and quality of information collected from interviews is greater. Unlike questionnaires, structured interviews do not require the respondent to give his/her answers in writing. This becomes more of an advantage as the motivation to write and writing skills of those being queried decreases. While some NCO's or enlisted personnel may be perfectly willing and able to answer a question in great detail during an interview, providing the same answer in writing may be another matter entirely. Thus, the interview is likely to provide more detail about each item of information sought.

Like the questionnaire, the structured interview insures that the same questions are asked each time the interview is conducted. This enables the recipient of the feedback to combine or tabulate responses for each item of interest. If the interview were not structured, as is characteristic of many interviews presently used for feedback, the tabulation and combination of responses would be difficult. Having each interviewee answer the same questions and combining the interviewees' responses to the questions allows the interviewer to determine whether a problem is widespread or merely applies to one or two persons.

Structured interviews allow the collector of feedback to gather information on events that he/she was not personally present to observe. As a result it tends to be a relatively efficient method of obtaining information on events that were not or could not be directly observed by the feedback collector.

While the structured interview has the questionnaire's advantage of providing the same kinds of information for each interviewee, it has flexibility that the questionnaire does not possess. During an interview the interviewee may indicate through his/her answer or in some other way that he/she does not understand the question being asked. When this occurs the interviewer can rephrase the question or otherwise clarify what is being asked. If the answer given by the interviewee is incomplete or raises additional questions, the interviewer can probe for further information. In this way, the interview can lead to unanticipated and perhaps more informative responses.

Interviews are generally better received by those who are being asked to furnish information than are questionnaires and other kinds of surveys. While this may seem to be a minor advantage, it is actually an important reason to use the interview method. The quantity and quality of the information gathered depends, in part, on the willingness of the persons surveyed to provide it. Individuals who are reluctant to complete a written survey may look favorably upon being interviewed. They often appreciate the opportunity to express their ideas openly to an interested, impartial listener.

The primary disadvantage of structured interviews as a method of gathering information is that the method requires a substantial investment in resources. Interviews require committing one or more skilled interviewers to the task over

an extended period of time. Interviews must be done on a one-to-one basis, with each interview lasting from 20 minutes to an hour. Because several persons at each level must typically be interviewed, the interview process can consume a considerable amount of time, depending on the size of the unit or organization being surveyed. In addition to the time spent conducting the interviews, reviewing and analyzing the information provided during interviews require many hours of a skilled analyst's time. Although data must be analyzed, no matter how they are collected, the data collected during interviews are more likely to consist of lengthy narration. Such narrative responses must be read, interpreted, and combined with similar responses for drawing conclusions, making the analysis of such data a difficult and tedious process.

The use of structured interviews requires a considerable investment in time prior to collecting the data for designing questions and planning the data gathering activities. Evaluators must often work hand-in-hand with training developers in order to be sure that the appropriate questions are asked in the most effective and efficient way. To design questions and plan the data collection, the evaluator/training developer must decide in advance what information is needed and who is most likely to be able to provide it. When the data gatherer has little or no idea of what he/she is looking for or has little time to plan the data gathering activities, the structured interview approach may not be the preferred method. In such cases, the data collector may need to initially use unstructured interviews or informal observation techniques to gather enough information to develop structured interviews for use at a later date.

Another potential problem with using structured interviews is that they require consistency and impartiality on the part of the interviewer. After asking the same questions in the same way a number of times, the interviewer may tend to become bored with the process and start to rephrase the questions, skip questions, press the interviewee for a quick answer, or unnecessarily prompt the interviewee in order to speed the interview. Unless the interviewer consistently asks the same questions in the same way, the quality of the information obtained will be adversely affected. Not only must the interviewer be consistent, but he/she must be careful to maintain impartiality throughout the interview. If the interviewer shows by his/her reactions that he/she wants a particular question answered in a particular way, then some interviewees may be so anxious to please the interviewer that they will be influenced to give the approved response. Although the interviewer often influences the interviewee unwittingly through a smile, nod, or by displaying additional interest, an unscrupulous interviewer will frequently lead the interviewee to give the desired answers to each question throughout the course of the interview.

Guidelines for Use

Despite some disadvantages associated with structured interviews, the method is highly recommended and should be included as a regular part of gathering information for feedback. Used correctly, this method can provide a great deal of detailed factual information and answer many different kinds of feedback

questions. Many of the guidelines listed for questionnaires are also applicable to structured interviews. Additional guidelines for using structured interviews for collecting feedback are presented below. Examples of interviews developed in accordance with these guidelines are included in Appendix B.

1. Interviews should be conducted with persons at different levels of the organization. Persons occupying jobs at the lower echelons should be interviewed as well as those holding leadership positions. Information is best obtained from those persons having first-hand knowledge of the situation. Those working at lower levels sometimes have information that is not available to their leaders. They may also provide a different perspective of events affecting the organization. Too often, only those occupying leadership positions are interviewed, which tends to present an incomplete and sometimes biased account of events.

2. The questions asked of interviewees should address areas that the interviewee is likely to have personal knowledge of, or experience with. Therefore, different questions may need to be developed for individuals at different levels or occupying different jobs. For example, the trainers in a unit might be asked about the training they conducted in the last two weeks, while those who are likely to have received training might be asked about the training that they received during the same period. Oddly enough the answers given by the trainers and the soldiers being trained do not always agree. This alone is reason enough to develop questions that can be asked of persons at different levels of the organization. Examples of questions asked at two different levels of a tank company are provided in Appendix B.

3. Interview questions should be directed toward identifying significant events or obtaining factual information. Questions asking soldiers for their opinions, attitudes, or evaluations should be kept to a minimum. For example, if you are interested in the training the soldiers are receiving in their unit, ask about the number of hours of training received in the last two weeks, the specific tasks that were trained, and whether soldiers actually got to perform the tasks during the training. Do not ask soldiers to rate their unit training on a five-point scale or to indicate how they feel about the training they receive in this unit. The former questions are specific and factually based; the latter are global and based on opinion. Specific questions call for specific information, while global questions generate very general and often useless information. Questions requiring soldiers to compare a large number of unspecified alternatives before selecting one or more of them are difficult to answer and should be avoided, especially when interview time is limited. For example, asking soldiers to identify the hardest tasks that they have to perform or asking them to identify the most critical tasks that they perform seldom yields much useful information.

4. Questions must be worded in a way that is meaningful to the soldiers. If terms used in the questions are intended to take on a particular meaning, the interviewer must define the terms so that the questions reflect the intended meaning. For example, when mechanics are asked how many hours they have spent in performing maintenance, their answer may refer to the number of hours they

were on duty or to the time they actually spent in troubleshooting or to wrench turning. If the interviewer is only interested in wrench turning, then the question may need to be phrased or explained in that way. Similarly, if a question inquires about the training conducted, the interviewer may need to describe the kinds of events that are considered to be training events.

5. As much as possible interviews should be conducted in a relatively private, quiet location where interruptions and other distractions can be kept to a minimum. In this way the flow of the interview is not impaired and the confidentiality of the interviewee's responses is maintained. However, a quiet location is not always available, and the interviewer is sometimes forced to conduct interviews at the soldier's work station in the presence of vehicle noise and other soldiers. When soldiers other than the interviewee are present, they tend to inject their comments into the interview, sometimes influencing the responses of the interviewee. When this occurs about all the interviewer can do is to politely ask the soldiers to refrain from commenting or suggest to those interrupting that they will get their chance to answer the questions later.

6. The persons conducting the interview should be intimately familiar with the questions to be asked and be prepared to clarify any questions that are not clear to the interviewees. On the other hand, the interviewer should stick to the script as much as possible, reading the questions as written, and deviating from the interview structure only when necessary. The interviewer must not show by his actions or words that he/she prefers one kind of answer over another. He/she must also avoid leading the interviewee to answer in a particular way. The interviewer must also record the interviewee's responses just as they are given, using the interviewee's words as much as possible. Care must be taken not to interpret the data or draw inferences from responses before they are recorded on the interview form. Once the data collection is complete, then inferences and conclusions can be drawn from the data.

7. To facilitate analysis of the information obtained, interview questions should be designed so that they elicit responses that can be coded and counted. This does not mean, however, that all questions must be answered by a number or a yes/no response. Of course, numerical and other countable responses make analysis much easier and quicker, and should be used when they provide the necessary information. Countable responses alone, however, may not proivde enough information to allow meaningful changes to be made. Narrative responses provide much of the detail and specificity that training developers need to improve their products. Caution must be used, however, not to place too much confidence in the narrative responses of any one interviewee. Typically an interview should elicit narrative as well as countable responses and should not rely entirely on either type of data.

Systematic Observation

An excellent, but seldom used, method of objectively obtaining feedback is through systematic observation. Systematic observation consists of methodically recording information about events as they occur, utilizing worksheets that list the events to be observed. This may be distinguished from casual observation in

which the observer is assigned to observe a particular activity and write down his/her observations or overall evaluation of it. Systematic observation requires the observer, at a minimum, to record his/her observations for each of the events listed on an observation worksheet. Additional comments may be included as needed to clarify recorded observations. The casual observer, on the other hand, usually operates under few, if any, constraints and is free to record or not record whatever observations he/she wishes.

Systematic observation may be used to gather feedback on training programs, the use of training materials, and for observing tests and evaluations of soldier performance. In combination with hands-on tests, observation is the preferred method for evaluating training programs. In addition to the guidelines presented at the end of this section, several good references providing detailed guidance in training program evaluation (see Appendix D) are available. These references may be very helpful in planning the observations and designing the necessary forms. In additon to evaluating formal training programs, observation can also be helpful in assessing the state of training or proficiency in units, especially as a check on information obtained through questionnaires and interviews. For example, specially designed observation worksheets can be used to spot check unit training or other unit activities such as maintenance. Examples of worksheets for spot checking training and maintenance activities in a tank company are included in Appendix C. To illustrate different worksheet formats, the training observation worksheet is presented in two different formats.

Positive and Negative Aspects

Systematic observation is perhaps the only method of obtaining feedback that provides training developers and evaluators with first-hand information about classroom and unit activities. In systematic observation, unbiased observers directly observe and record training events or other unit activities as they occur. In this way the information does not pass through the selective filters of memory and self interest that may distort information acquired through interviews and questionnaires. Thus the information obtained through systematic observation is generally more objective and accurate than questionnaire and interview data, assuming that the observers are objective and unbiased.

In systematic observation, the observer knows in advance what observations he/she will make. He/she has been trained to record the data accurately and objectively. Because the observer knows what he/she is looking for and records it accurately, the chances are improved that the needed information will be obtained. Many things that might go unnoticed by the casual observer will be picked up and recorded by the trained observer, particularly when aided by the use of structured observation worksheets. Aside from the information recorded on the observation worksheets, direct observation of the problems associated with various unit activities may suggest the causes of the problems and solutions to them. Insight may also be obtained regarding which changes are feasible and likely to be most beneficial.

Systematic observation provides an alternative to more subjective methods for evaluating training programs. Observational methods can be used to determine

if the training is conducted in accordance with the training plan, and to document the occurrence of unplanned events during training. They can also be used to provide accurate information about the training environment. Tests can be observed to insure that the tests are conducted and scored properly and that test results are not contaminated by poor test administration or biased scoring. Whereas questioning trainees, instructors, and examiners may in theory provide essentially the same information about the training program, confidence in the validity of the information obtained is increased substantially by direct observations by unbiased observers of the activity being evaluated.

The biggest drawback to using systematic observation is that the method can be very time-consuming. When used for spot checking unit training or maintenance activities for the purpose of verifying interview or questionnaire data, observational data can be collected in a relatively short time frame. However, if systematic observation is the primary method of obtaining feedback, such as might be the case in evaluating a training program, the time invested by one or more observers may be considerable. The time necessary to compile and analyze observational data can also be considerable, particularly when the data contain a number of written comments. In defense of systematic observation, it should be noted that the time spent in observing unit activities is time well spent. The loss of efficiency is more than compensated for by the gain in the accuracy of the information obtained.

Another disadvantage of using systematic observation is the requirement for trained observers. Although an observer need not be an expert on the subjects trained or other activities that are being observed, he/she must be sufficiently familiar with the activities being observed to intelligently make the required observations. More importantly, the observer must be trained to be an objective observer who faithfully records what he/she sees. Generally observers must be familiarized with each item on the observation worksheet and given practice in completing the worksheets prior to collecting any feedback data. They will also need to be trained to distinguish between what they observe and their interpretation of what they observe. Such training is absolutely essential to the success of systematic observation as a method of gathering feedback.

Even with training, not everyone makes a good observer. Some observers lack the ability to adequately express their observations in writing. If asked, they can explain verbally what they saw, but are unable to record their observations as written comments on the worksheet. Other observers, who possess the necessary writing skills, write their interpretations of what they saw or their conclusions rather than simply recording what they actually observed. Still others lack the discipline required to perform systematic observation. On-site observation can be uncomfortable, boring, and difficult. It requires the observer to make specified observations of certain events repeatedly, usually according to a schedule and often outside during inclement weather. Self-disciplined observers are therefore a necessity.

Systematic observation requires careful planning and coordination with the group being observed. It involves determining in advance what events are going to be observed, and informing the group that an observer will be collecting data

on certain events within a specified time frame. It involves coordinating with the observed unit to insure that the observer has free access to the sites and events of interest. It requires knowing when and where unit activities are occurring, so that an observer will be present to cover important events.

Observers are not always greeted by those being observed with open arms, particularly when the observers are perceived as evaluators who are there to judge the unit or the trainers. Sometimes those observed complain that the observers are distracting and interfere with the activities being observed, thus preventing them from doing their jobs as they see fit. The complaint has some merit; the presence of an outside observer cannot help but be somewhat distracting and may affect the conduct of the activities being observed. The extent to which events are conducted differently than they would be in the observer's absence adversely affects the accuracy of the feedback gathered. But the effect on the event observed may be positive if the event is improved because it was observed. Occasionally, attempts will be made to bar observers from certain kinds of events. The reason given for excluding the observers may vary, but the result is the same -- less accurate feedback.

Guidelines for Use

Most major feedback collection efforts should employ systematic observation as one of the methods for collecting feedback. The tendency to rely solely on subjective data may be reduced through the judicial application of observational techniques. As a check on the data provided by other methods and for a number of other purposes, systematic observation is an under-utilized tool that has great potential for increasing the quantity and quality of information available for feedback. To realize the full potential of systematic observation, forethought, planning, and careful execution of the method are required. Specific guidelines for using systematic observation for collecting feedback are provided below.

1. A form or worksheet should be designed to be used in making the observations. The worksheet should contain a listing of the items or events to be observed. The items should list events as they might be expected to occur under ordinary circumstances. The kinds of items included on the worksheet will of course depend upon the particular activities being observed. Item listings for various activities are included in the references under "Observation" in Appendix D. Examples of items useful in evaluating training and maintenance activities are shown in Appendix C. Other items may be developed for specific applications, but all items must meet certain minimum requirements. The items should deal with things or activities that can be directly observed. Things should be described in terms of observable properties, and activities should be described in terms of observable behaviors. To facilitate the recording of observations, the items should be stated so that they can be answered Yes or No, with additional space provided for more detailed comments. Examples of worksheets that follow these guidelines are included in Appendix C. The formatting of the worksheets included in this appendix are not sacred; they can be modified as needed to make it easier to record observations. However, we must advise against making any changes that make it necessary for observers to record their subjective

judgments or evaluations of events. Avoid items that require the observer to rate each activity or to judge aspects of performance or the environment as satisfactory or unsatisfactory. Observers should use comment space for providing additional details about what they observed. This space must not be used for recording subjective judgments or opinions about personalities, the unit, or unit activities. That is, observers should record specific behavior and facts that were observed and not their interpretation of events. For example, they shouldn't record a general judgment that "the exercise controller didn't know what he was doing," but should instead record specific problems that were noted, such as scoring discrepancies or violations of SOP's. The items on each worksheet should be preceded by a heading. The heading should identify the activity being observed, when and where the activity occurred, and who was present to include the observer.

2. All forms or worksheets developed for observing training, other activities or the environment should be pilot tested prior to using them to collect information for feedback purposes. The conditions under which the observation forms are tested should resemble as closely as possible those under which the forms will eventually be used. Typical observers should be among those testing the forms. The events observed, whether they be formal training events or day-to-day unit activities should also be representative of the kinds of events on which observational feedback will be collected. New observation forms may be piloted in conjunction with the training that each observer must undergo to certify that observers are knowledgeable about observational methods and the observation forms. By piloting the forms, it can be determined which items produce useful feedback and which observers are most proficient in gathering the feedback.

3. If more than one or two observers will be involved in collecting the feedback, it may be advisable to conduct a workshop to train the observers. The observers should be told the purpose of the workshop, i.e., to train them to use the observation worksheets for collecting information to be used as feedback. The importance of learning how to accurately record their observations should be stressed. The instructor should go over each item to be included on the worksheet with the group to insure that they will know what it is they are to look for, how they will know when it happens or does not happen, and how they should record their observations. The instructor should get feedback from the observers to insure that they understand the items in the same way, so that there will be standardization across observers. If a filmed or televised instructional sequence is available, the instructor can run the tape or film, commenting on those parts of the film that relate to the items on the worksheet. A necessary part of the workshop involves giving the observers practice in using the worksheets. The same events should be observed by all observers so that the observations can later be compared. The events observed should be actual events that resemble those on which feedback will later be collected. The information gathered during this practice session should be reviewed in class, and the worksheets completed by the observers should be collected by the instructor, so that the instructor can determine which trainees record their observations most accurately and completely. References listed under "Observation" in Appendix D provide additional guidance in selecting and training observers.

4. Those in charge of collecting the feedback should assure the unit or the trainers that the purpose of the observations is not to evaluate them, but to assess the training program, task proficiency, or the usefulness of training materials. Training developers and evaluators should be open and honest with trainers or unit personnel about their purpose for being there. Observers should not attempt to hide recorded notes or completed worksheets from trainers or unit personnel. Observers should be cautioned not to interfere with the conduct of training, testing or other unit activities, and should not interact with trainers or unit personnel any more than is essential to collect the necessary observations. At the same time observers must be told the importance of remaining close enough to the activities being observed to insure that the data gathered accurately reflect what occurred.

5. Observers must be carefully selected on the basis of their performance during the workshop. Completed worksheets should be reviewed in order to determine each observer's proficiency in recording the required observations. The best observers are those who record what they observe in detail on the observation worksheets, with a minimum of interpretation. Those who jump to conclusions or make sweeping generalizations on the basis of what they saw do not make good observers.

6. Observers must go where the action is. To provide accurate feedback, they must observe important training events or unit activities first hand. If the activities of interest are occurring in the motor pool or in a unit training area, then that's where the observer must be. Arrangement should be made with trainers or unit leaders to allow observers free access to various sites. Observers should have the option of appearing at these sites unannounced. However, observers should take care not to disrupt unit activities by unannounced appearances at sites where they are not known. Unit personnel must be informed ahead of time that observers will be in their area from time to time, and should know generally what the observers will be doing. It may not be realistic for observers to arrive completely unannounced, but it is usually not necessary to announce when they will be where or the exact activities that will be observed. For scheduled events, such as scheduled training, observers should arrive in the area at least 10 minutes before the event is scheduled to begin. When the same event is occurring simultaneously at several stations or locations, additional observers will be needed to provide adequate coverage of the event.

7. Try to minimize the boredom and discomfort experienced by observers. In planning data collection activities keep the observers' interests in mind. Avoid over-committing observers by limiting the number of data elements that they are required to collect at any one time and the hours spent in observation activities on any one day. In addition provisions should be made for the well-being and comfort of the observers. Such provisions might include transportation to and from the site, messing facilities near the site and the opportunity to use them, and a temporary shelter where observers can go from time to time to escape the elements. This will result in better motivated observers, and hence, more accurate and complete feedback.

Testing

As feedback, test results may be used to provide objective information about the skills and knowledges acquired during a particular training program or to assess the proficiency of soldiers in the field in performing certain tasks. Training developers use testing to determine if the training programs, training support materials, and on-the-job experiences result in the soldiers being proficient on specific tasks. Two types of test results may be useful as feedback: (1) Test results from tests that are a part of a training program or unit testing program (i.e., results from existing tests or existing records); and (2) Test results from tests specially designed for feedback purposes. Because evaluators and training developers are primarily interested (or should be interested) in performance on specific tasks under a given set of conditions and standards, most test results used for feedback consist of records of hands-on performances. To the extent that hands-on tests are not included as a normal part of institutional and unit testing, feedback collectors will need to design special tests for obtaining feedback on task performance. Considerations involved in using results of existing tests or designing special tests to gather feedback will be discussed further in the guidelines for use of testing.

Positive and Negative Aspects

As a source of feedback, test results are probably more highly regarded by training developers and evaluators than any other type of information. When asked how much confidence they place in the accuracy and reliability of various types of feedback they collect, training developers and evaluators typically reserve their highest marks for test results. This high degree of confidence may be attributed to the perception of testing as a rigorous objective method for assessing performance. While this perception probably is accurate when tests are conducted properly, the confidence is misplaced for poorly administered tests.

Some common problems that may affect the accuracy and reliability of tests include: testing the wrong tasks; vaguely stated test instructions; tasks tested using inappropriate or unrealistic standards, and under wrong conditions; and erroneous test results due to biased or inconsistent scoring among examiners. Because test administration is rarely observed, such problems often go undetected and questionable test results are accepted as accurate indices of performance.

When tests are conducted properly by unbiased examiners, they are among the more objective methods for obtaining feedback. The objectivity of tests makes them acceptable measures of training effectiveness and permits them to be used as criteria against which other sources of feedback (e.g., estimates of task proficiency, questionnaire responses, and interview results) are compared. Tests administered before and after participation in unit training activities or institutional training programs may be used as one index of the effectiveness of that training. For checking on the accuracy or various estimates of task proficiency, no better criterion exists for comparison purposes than hands-on tests of performance.

Testing, particularly hands-on testing, has not been the preferred choice of training developers and evaluators for collecting performance information. The reasons for this are many. Foremost among these is the heavy demand placed on the resources of trainers, units, and feedback collectors alike by hands-on tests. All feedback methods require the expenditure of time and resources by feedback collectors, and probably by those from whom the feedback is collected. Of the various methods, hands-on testing requires the highest expenditure of resources by collectors and suppliers of feedback. A resource that testing consumes in large quantities is time -- a resource that is always in short supply. In terms of personnel resources, hands-on tests require a sample of individuals to be tested, one or more unbiased, trained examiners, someone to make arrangements for provision of the needed facilities and equipment, and possibly observers to monitor the administration of the tests. Other resources required for testing are carefully designed tests and scoresheets, adequate testing facilities, and equipment to be used in the testing process. The large resource requirements of hands-on tests probably explain why they are not a larger part of the feedback picture than they are presently. The resources required in testing depend to a considerable degree on the number of individuals or groups to be tested. Feedback gatherers often assume that everyone must be tested if test results are to be useful for feedback purposes. While some applications of feedback do require that most or all of the individuals participating in an activity be tested (e.g., in evaluating training effectiveness), testing a small sample of individuals can sometimes provide sufficient information. To discover which tasks are performed well by a unit, a representative sample of unit personnel can be tested on a set of tasks. For spot-checking the accuracy of data obtained through the use of other methods, testing a small sample from the group from whom feedback is being sought is often sufficient.

Guidelines for Use

1. If the feedback desired consists of information about proficiency in performing specific tasks hands-on, then the tests should require hands-on performance of these tasks. The tasks selected for the test should be those that the soldiers tested might reasonably be expected to perform on the basis of their training and experience. The tasks tested should be stated in terms of measurable performances, with conditions and standards that are consistent with those specified in the training objectives or in the Soldiers Manual. Test instructions should clearly state what the soldiers are expected to do and should be presented in a standardized manner such that each soldier receives the same instructions. Score sheets should be designed so that the examiner can record performance efficiently and objectively. Distinct parts of the task that can be readily observed should be listed on the score sheet, so that examiners can record, in addition to an overall evaluation, those parts of the task that were or were not performed correctly. Scoring each distinct observable part of a task provides additional information about that task which may be used by training developers or evaluators to pinpoint the reasons for poor task performance. There are of course practical limitations on the amount of detail that an examiner can record for any given task. But efforts should be made to record as much detail as possible when the data are to be used as feedback.

2. Testing requires the full cooperation of trainers in the unit or at the institution. Arrangements must be made well in advance to secure the needed equipment, facilities, and personnel for the period during which the testing is to be conducted. Testing is an activity that requires careful planning and execution by those conducting the tests. Because the planning and execution typically involve the trainers in some capacity, it is essential to enlist their cooperation. Cooperation of unit and institutional trainers is more likely if the test results are not perceived as a threat. If institutional and unit trainers think that test results collected will be used against them, they may refuse to permit training developers or evaluators to conduct the tests. One way to prevent test results from being used inappropriately is to maintain the anonymity of individuals and units in reporting results. Anonymity becomes more important as the level at which the results are reported increases. Because training management is accomplished at the battalion level, generally there is no reason to report results to levels above this level. When reporting results to the battalion commander, it is recommended that the companies not be identified in reporting the results. Similarly platoons should not be identified in reporting results to the company commander. Even at lower levels, it may not be advisable to report individual test results by name. While obtaining the cooperation of the unit may entail sharing the test results with unit leaders, it is important not to alienate individuals by identifying those who performed poorly on the tests. When results are to be reported outside of the battalion, or used by feedback collectors, all references to individuals and units should be omitted.

3. Proper test administration is essential to obtaining accurate test results. The tests should be administered by independent, unbiased observers whenever possible. This presents no problem when the tests are developed and administered by training developers and evaluators whose primary purpose is to obtain feedback. However, when tests are administered by those who conduct the training at the institution or in the unit, there may be a tendency towards leniency in conducting and scoring the tests. When tests are <u>not</u> administered by independent examiners, trained observers from the evaluation team should observe the testing to insure that the tests are administered properly. A form for making test observations is included in Appendix C. Additional guidance in observation of testing may be found in references included under "Observation" and "Testing" in Appendix D.

4. Performance tests should be used to check on the accuracy of results obtained through less objective methods. Soldiers and their supervisors are often asked to provide estimates of how well they can perform specific tasks. Several precautions should be taken in obtaining such proficiency estimates. The performance being appraised should be stated as precisely and completely as possible. Rating scales should be phrased in terms of explicit observable measures of performance rather than in general terms such as "average", "below average", etc. Raters should be trained and have had experience with the rating scales and tasks being rated. As a further measure, the accuracy of these estimates should be checked by selecting a sample of the tasks and testing, hands-on, the performance of individual soldiers on the tasks. At the time the proficiency estimates are obtained, the soldiers and their supervisors

should know they will be tested on selected tasks, but they should not be told exactly which tasks are to be selected. Unless proficiency estimates are shown to be relatively good predictors of hands-on test performance, such estimates should not be used for feedback purposes.

SECTION III - INTEGRATION OF METHODS

Now that each available feedback collection method has been discussed, this section addresses various issues related to the planning and conduct of feedback activities. The gathering of feedback is a flexible process that must be adapted to the requirements of particular situations. None of the feedback methods discussed in this chapter is sufficient in and of itself for gathering feedback in all situations which may arise. A systematic approach is thus needed in which methods are integrated and tailored to meet the needs at hand. There is no set formula or procedure through which an appropriate mix of methods can be selected and applied. There are general approaches, such as the sequential mix mentioned earlier in this paper. In this approach, initial indications of problems are obtained through informal feedback or review of existing performance records. The scope of problems identified is then determined through the administration of questionnaires, and further details are obtained during follow-up interviews. If further information appears to be needed at this point, it may be obtained through administration of performance tests or conduct of structured observations using precoded worksheets indicating what is to be observed. While such a general approach has the advantage of leading to the collection of more precise data as needed, it may not be applicable in all situations. Resource and time constraints may not allow the sequential application of all available methods. Informal feedback and existing records should not be solely relied upon for initial indications of problem areas, since problems may surface during interviews or observations. Feedback gathering should move toward the collection of more detailed and precise data as resources allow, but there is no set sequence for the application of methods. The selection of feedback methods should be based upon prescribed criteria, and some of these are delineated below. This is followed by discussion of other general issues, including sampling considerations and the structure of feedback collection efforts.

Criteria for Method Selection

There are numerous factors which interdependently impact upon the choice of feedback methods and the design of a feedback collection exercise. Among these are the resources available, the established accuracy of the methods available, and the scope of the area being addressed. These factors should be jointly considered as criteria for making decisions on how to collect feedback.

The appropriate resources for application of each feedback method have been addressed in previous sections, so only general resource guidelines will be offered here. Most DTD's and DOES's do not have sufficient personnel resources to support large-scale, continuous collection of feedback. The acquisition of feedback must thus be efficient in order to be doable within available resources. But the feedback process should not be short-changed; a minimum amount of resources is needed to collect adequate feedback. Training developers and evaluators cannot collect adequate feedback while sitting in their offices receiving sporadic informal feedback and mailing questionnaires. As discussed earlier, informal feedback is usually very general and mailed questionnaires are frequently not returned. Training developers and evaluators

must have the resources available to collect feedback first-hand by visiting training sites in the institution and in the field and by interacting directly with users of products of the training development process. Training developers should thus not try to collect feedback using just their own limited resources, but they should work through and interact with Branch Training Teams (BTT's) and other training evaluation teams. The combination of DTD and DOES resources may help eliminate problems in resource shortages, so it should be actively encouraged. If such pooling does not provide adequate resources for collecting needed feedback, then the scope of the area being addressed or the questions being asked should be reduced. For example, feedback may be needed on the performance of all tasks within a given MOS, but resources may not be available to support the collection of such extensive data at one time. Managers of feedback efforts should acknowledge the existence of such a situation and tailor their efforts to collect specific data on a sample of the tasks of interest rather than collecting general data on all tasks. Precise data on a subset of tasks is likely to be more valuable and useful than general and perhaps inaccurate data on a wide range of tasks. This issue is further addressed below in a discussion of the accuracy and scope of feedback approaches.

Resource concerns relate not only to the number of personnel involved in feedback collection but also to their qualifications and background. The ideal feedback collector should have experience in both the subject matter being addressed and in the collection methods being used. Subject matter experts do not automatically make good feedback collectors. Knowledge of the subject matter being addressed is helpful, particularly in the gathering of specific details. This is one of the reasons that training developers should be included on feedback collection teams. But expert knowledge can sometimes be a hindrance. For example, during structured observations subject matter experts may become so intensely involved in the information being presented that they miss seeing major problems in how it is presented. Feedback collectors should be given training and practice in the application of the methods being used, especially interviews and observations. They must learn to record the data as presented, and not their interpretations of the data. A workshop in which the actual data collection forms are discussed and used is the best way to train data collectors. Guidance for conduct of such a workshop is included in the references at Appendix D.

As pointed out above, it is very important that the feedback process be efficient. But one consideration that is even more important is that the data gathered as feedback be accurate. Since major revisions in training programs and training support materials are based upon feedback, this information must accurately represent the needs and experiences of users. Concerns for accuracy should dictate the ways in which available resources are utilized. Subjective data having unknown accuracy should always be checked against relatively objective data. If a questionnaire or interview is used which has not previously been validated, then it should be supplemented by structured observations or tests. Validation of a questionnaire or other data collection form involves more than just the administration of the form to a sample of personnel and the noting and correction of general problems in wording or format. A form is validated when the responses gathered using it have been shown to agree with

data gathered using more objective criteria. Since few if any feedback collection forms have been validated, all feedback efforts should employ objective measures such as structured observations and testing at least to a limited extent. It may not be necessary to check the accuracy of every subjective response obtained, but spot checks should always be conducted. If such checks indicate problems with the accuracy of questionnaire or interview responses then these responses should be disregarded and replaced with more objective data. All available methods must thus be integrated in a feedback collection effort, with the degree to which each method is included in the mix depending upon the demonstrated accuracy of the data obtained. Sufficient resources must be committed to insure the collection of accurate feedback. If this is not done, inappropriate revisions may be made in training, and the feedback can do more harm than good. When resources are limited, the scope of the effort should be reduced rather than the precision of the data. Absolute accuracy of feedback can never be achieved, but the highest possible level of accuracy should always be sought.

The discussion above has indicated ways in which considerations of scope interact with other factors impacting upon feedback method selection. Resource and accuracy concerns may limit the scope of a feedback collection effort. Inversely, the scope of an effort influences resource and accuracy considerations. If feedback is needed on one small segment of a training program, it can probably be collected by one person within the Center/School without detailed accuracy checks. But if feedback is needed on major portions of a training program, it must be collected by a team of training developers and evaluators visiting the field and gathering various types of data as accuracy checks. The scope of the effort must be delimited during planning of feedback collection, and resources must be made available which realistically allow data collection of that scope to be accomplished. Otherwise, the quality of the data collected will likely prove to be unacceptable.

Sampling

Another aspect of feedback collection related to the factors discussed above is sampling considerations. It is generally not appropriate to base feedback on just one source, as is sometimes done in informal feedback collection. But it is often not possible to gather feedback from all users of a training product or on all aspects of a training program, due to resource and time limitations. After selecting the appropriate methods to be used, training developers and evaluators must then select the appropriate samples to which they are to be applied. Sampling relates not only to the specification of individuals and units from whom information is to be collected, but also to the specification of tasks and areas of concern which are to be addressed. In many cases sampling can be made more efficient by integrating these two dimensions. That is, different samples of individuals can provide feedback on different sets of tasks or areas of concern. Particular sources may provide only part of the information needed, but the overall data collection effort should provide all the information needed.

AH

Sampling is not usually a concern in the gathering of informal feedback, since such feedback is generally unsolicited and is passively received by training developers and evaluators. But sampling considerations are important with all other feedback methods, especially those that are relatively resource-intensive, such as interviews, observations, and testing. Sampling of feedback is basically a balancing act between obtaining enough data, but not too much. Sufficient data must be obtained to accurately represent the overall situation, but resources should not be consumed by collection of large amounts of data which may be redundant and difficult to manage. Statistical sampling procedures often cannot be applied in feedback collection, since real-world constraints often preclude truly random sampling. It is thus difficult to specify the size of the sample needed in each feedback collection effort. For this reason, only general sampling guidelines are offered below. But this does not mean that training developers and evaluators should not be concerned with statistical sampling procedures. Sampling should always be conducted as carefully and rigorously as the situation allows. Statisticians and sampling textbooks should be consulted to supplement the guidelines offered here.

To support overall evaluations of training programs and support materials, feedback should be gathered from more than one unit or a small set of units in one geographic area. Contingencies such as range and ammunition availability vary in different locations, and training programs are often implemented somewhat differently in different units. To obtain an overall picture of the situation in the field, feedback should thus be sampled from several units in each of several geographic locations. Feedback obtained within a unit should be gathered from all levels of personnel involved in the training process. For example, if a questionnaire is administered within a battalion on the utility of particular training materials, it should be administered not just to the senior leadership, but to all levels of users. Feedback on products of the training development process should be obtained from personnel directly using those products, including NCO's serving as trainers and soldiers receiving training. When individuals are asked to provide feedback through a questionnaire or interview, they should not be asked to provide so many responses that they become fatigued. When respondents become fatigued, the quality of their responses is likely to suffer. Questions should be sampled so that no individual is required to participate in a survey or interview for more than one hour. As suggested above, different samples of questions can be addressed by different individuals in order to provide the complete feedback needed.

Sometimes it is practically impossible to specify the size of the sample needed before embarking on a feedback collection effort. In these cases a useful strategy is to perform preliminary analyses of the data while they are being collected and to stop collecting further data when they become redundant and provide no new information. Collectors of feedback should not expect to obtain every piece of relevant data that may be available. But they should be aware of general sampling considerations and should sample sufficient data to be representative of the situation existing in institutional or unit training. Decisions on sample sizes needed are best based on experience with feedback methods and with the types of data collected.

Structure of Feedback Collection

In this section the information in previous sections is summarized in terms of the general structure needed in feedback collection efforts. The first point that should be made is that feedback collection should be planned in advance and structured as highly as possible. Specific uses should be planned for all data that are to be collected, and sufficient copies of structured forms should be available for recording of each type of data. The resources needed should be organized, and the samples from which data are to be collected should be clearly identified. All feedback collection forms should be thoroughly pilot tested, and all data collectors should be trained in the use of these forms.

Another key point is that it is very difficult to over-coordinate the collection of feedback. Thorough prior coordination will eliminate many problems that commonly occur during data collection. Training developers and evaluators should continually coordinate their efforts to avoid duplications and to make certain that the needed data are obtained. Coordination with those units or individuals providing the feedback will help insure that structured data collection schedules can be adhered to. Feedback collectors should enter units looking for specific events or results, rather than looking for whatever happens to be going on. This requires careful coordination with unit personnel, since unit training schedules are highly dynamic. Coordination with the providers of feedback should not stop once the data are collected, but should continue as long as useful information is being transmitted. Feedback providers should be informed as to what conclusions were reached from the data they provided and what resulting actions took place. They can then provide feedback on the effectiveness of these actions, and a continuous cycle of feedback can be established.

SECTION IV - MANAGING FEEDBACK

Feedback collected by any means is valuable only when it can be used to make informed decisions which lead to needed changes in training programs or materials. In addition to the level of detail and accuracy of feedback, the form in which the data are available and the speed and ease with which they can be accessed are important determinants of how feedback will be used. Rapid access to data that are in a readily usable form is dependent upon the installation of an integrated data management system. Such a system is essential for organizations such as DTD's and DOES's who collect and use feedback on a regular basis. Differences among these organizations in the kinds of feedback available and the ways in which they are used make it impossible to specify in detail the characteristics needed in a data management system. But systems for managing feedback should always incorporate certain general features in their design and follow certain guiding principles. These features and principles are discussed in the remainder of this section.

Automation of Data Management

A primary principle for guiding data management is that it should be automated to the maximum extent possible. Although computer support is not necessarily required for effective data management, the increased efficiency and flexibility provided by computers strongly recommends their use. The increased efficiency results from computers' ability to analyze large volumes of data quickly and to provide instantaneous access to data of interest. Because computers can be programmed to quickly reorganize information in different forms (e.g., graphs, tables, etc.) and to extract and compare any data of interest, computers provide training developers and evaluators with the flexibility needed to get the maximum benefit from the data they collect.

The advantages of automated data management far outweigh any arguments that might be raised against the use of computers for that purpose. In the past, the use of computers for widespread applications in the management of feedback may have been rightfully considered as too costly. Now it is not only cost-effective to use computers for such applications, but it may actually be more costly not to use them. An elaborate computer system will not be required to manage feedback at the typical Center/School; minicomputers are available which can accomplish this task with the necessary supporting equipment. A minicomputer with expanded data storage space, peripheral input/output terminals, and telecommunications support for interaction with other automated data systems should meet the feedback management needs of any Center/School. Such systems should be standardized to some extent across installations to ease the transfer of data among them, but they should be flexible enough to meet the needs of each DTD and DOES.

Because computers allow easy access to data, they sometimes present security problems. Such problems are not unique to computers, but occur whenever large volumes of data are maintained, some of which are classified or of a sensitive nature. Special techniques are available for securing information stored in computers, but the need for implementation of such techniques can be eliminated

by avoiding the use of sensitive or classified information as feedback. Information on task proficiencies, training effectiveness, and usefulness of training materials is generally not classified, and it is usually not considered sensitive when unit and individual identifiers have been removed. These types of information comprise the bulk of data that are used for feedback purposes. The benefits gained from the use of sensitive information for feedback purposes are generally not worth the additional workload required to maintain it. For this reason, it is recommended that the entry of classified or sensitive information into a feedback data base be minimized, if not completely avoided. As suggested earlier, the sources of feedback should be kept anonymous in order to reduce sensitivity problems with the data.

Preferably a computer dedicated to the maintenance and analysis of feedback should be located on-site at each DOES. DOES personnel would be responsible for determining what information is included in the data base, but the data stored in the computer should be readily accessible by DTD personnel and other training developers. Remote terminals linked with the computer should be located at DTD and other appropriate locations, permitting direct access to feedback by training developers. The data stored in the computer should be limited to feedback that is known to be accurate and has a clearly defined use. Care should be taken to insure that the computer does not become a depository for whatever large pools of unvalidated subjective data that are available. DOES personnel should maintain knowledge of the methods used to collect various types of available data and should filter out data which have not been shown to be accurate or which do not have discernible utility.

Many of the characteristics needed in a feedback management system may be found in existing data management and analysis packages. These packages may be used, as is, or adapted to fit the particular requirements of each Center/School. If existing packages are not totally adequate, then prototype data analysis and management techniques should be developed based upon existing data and expanded as more data become available. In selecting or developing a data management system, close attention should be given to how the data are organized for retrieval from the data base. The system should also provide some means for analyzing the data and arriving at decisions on the basis of the analysis. A capability should also be available for presenting the data in alternative easy-to-use formats. The data management system should be dynamic so that the data base can be updated as new data become available. It should also provide a mechanism whereby changes to training programs and materials based on feedback are recorded and tracked to determine their effects. Finally, a data management system used for feedback should make provision for integrating its data base with relevant data from other systems and Centers/Schools. These needs are discussed in further detail below.

Organization of Data

A primary advantage of computer-based data management systems over manual systems is the capability to retrieve needed data quickly. The facility with which the automated system performs this function depends upon the manner in which the data are organized for retrieval. Feedback relating to training

materials can be organized according to the particular materials to which the feedback applies. Feedback on proficiency can be organized by task and/or mission, and feedback on the effectiveness of training programs can be organized by class or block of instruction. By identifying the publication, tasks, or class for which information is needed, the feedback user should be able to find the relevant data. In order to pinpoint feedback more precisely, the system should permit the user to specify other parameters. For example, in order to allow retrieval of information gathered during a given time frame, the stored data must be identified according to when they were collected. Similarly, if data for different types of units (e.g., active versus reserve units) are to be analyzed separately, then data must be tagged as to the type of unit from which they were collected. There may also be a need to index data so that only soldiers having certain experience or training levels contribute data to particular analyses. It may thus be necessary to include demographic information on the sources of feedback in the data base.

The number and types of parameters on which the data are indexed depends upon the needs of the training developers and evaluators involved. When data are retrieved from the data base, the output should include all relevant identifiers that apply. In addition to the identifiers described above (e.g., time frame, type of unit), the output should include information about the size and composition of the sample from which the data were obtained. The instrument or method used in collecting the data or the file from which the data were retrieved should also be listed or briefly described in the output. The inclusion of identifiers such as these is necessary when data are obtained from diverse sources using a variety of data collection methods and instruments. Identifiers should not be included to the extent that they take up more storage space than the data themselves, but sufficient information should be stored to allow the organization and retrieval of feedback along several dimensions. In this way, users of feedback will have the flexibility to select the precise information that they need.

Integration of Data

As suggested earlier, the methods used to collect feedback should be integrated within a total system. Similarly, the data collected should be integrated. Often one method is used for checking the accuracy of data produced by another method. Generally, when two methods aimed at obtaining similar information produce largely discrepant results, the data produced by the more objective method are given more credence in deciding which changes should be made. When different methods yield essentially the same results, then more confidence can be placed in the data and resulting conclusions than if they were produced by a single methodology. Usually, data collected by different methods are in somewhat different forms and cannot directly be combined for analysis purposes, even when they are mutually supportive. But when there is more than one source of feedback about a particular training program or document, it is advisable to examine and compare data from all sources. Often one source will complement the information provided by another source, even if the data cannot be combined directly. For example, results of a test given at the end of a block of instruction may indicate that the training was not effective, but give little indication as to

why. Observations made during the instruction may pinpoint the reason for poor performance as being too little practice on several of the tasks. To identify the problem and the correctable reason for it, both sources of information are needed. In similar ways, data produced by questionnaires, interviews, observations, and tests may all complement each other.

Data collected from different groups by the same method may also complement each other. For example, information collected from personnel at lower echelons may reinforce and expand upon that collected from personnel at higher levels. When information is gathered from more than one unit of the same type within a relatively short period of time, the data may be combined during analysis and conclusions may be extended to the units as a group. When units are selected in a representative fashion, conclusions may be generalized to other units of that type. Training developers and evaluators may at times collect the same information in the same general way without being aware of the duplication. Such duplications should be identified and the data should be combined so that training developers and evaluators can benefit from all available information. For certain kinds of feedback, integrating the data with those gathered by other Centers/Schools may lead to the identification of problems existing across branches or even across the entire Army. Integrating data in this way would require extensive communications among computers located at different sites and a greater degree of standardization of data collection and management procedures than presently exists.

Regardless of whether the data are gathered by different methods, from different groups, or by different organizations, the important point to remember is that related data should be integrated to the maximum extent possible. Information in the data base should be catalogued and cross-referenced with related available data. Thus, if both observational and test data are available in the data base for a particular block of instruction, any attempt to access one type of data will automatically remind the user of other related data that are available. The manner in and extent to which data are combined during storage and analysis will vary with the types of information involved, but related data should always be compared and combined to the maximum extent possible. Such integration will generally increase the amount of information available to training developers and evaluators, and this will lead to more effective decision-making based upon feedback.

Data Analysis and Decision-Making

The methods used for analyzing data depend on the kinds of data collected and the form in which they are obtained. Responses to questionnaires or interviews must usually be coded so that each response is assigned to one of a small number of categories before being entered into a computer for analysis. For example, if soldiers are asked what activities other than training they participated in during the last two weeks, the responses given may vary widely. In order to facilitate the analysis and reporting of the data, the various responses could be assigned to categories such as the following: (1) maintenance, (2) post detail, (3) inspections, and (4) other activities. Sometimes the response categories are known in advance and can be preprinted on the data

collection form as response alternatives. In other cases it may be necessary to derive the categories after the data have been collected. The basic data in both cases will consist of frequency counts of the number of responses given in each of the categories.

As far as statistical analysis of feedback is concerned, the best general guidance that can be given is to keep the analysis as simple and clearly understandable as possible. The purpose of analyzing feedback is to reach and support conclusions and recommendations about changes needed in training programs and materials. Many of the personnel involved in making decisions about these recommendations do not have an extensive background in data analysis; it will thus be necessary to keep the presentations of analyses simple and straightforward so that these individuals can see how the data support the conclusions reached. Also, the application of sophisticated analysis techniques frequently involves numerous assumptions about the characteristics of the data being analyzed. Data gathered in an operational military environment often do not meet many of these assumptions. It is thus best to cautiously apply sophisticated analysis techniques in the processing of feedback, and to consult a statistician or reference text to make sure that the technqies are applied appropriately. The analysis of feedback may be limited to the descriptive level, in terms of frequency counts, percentages, or proportions. For example, the data of interest may be the percentage of soldiers failing to perform a given task to standard or the proportion of those responding "yes" to a particular question. Summary statistics such as means or medians can be used to concisely describe data collected, but in some cases use of such statistics may not be appropriate. For example, two sets of data ranging along a five-point scale may have the same mean but considerably different distributions. It is thus prudent to plot the data graphically or otherwise examine the distributions of data obtained, as well as calculate summary statistics. In analysis of feedback, one should stick as closely to the original form of the data as possible and take care to insure that any summary statistics used are appropriate.

Techniques of inferential statistics should be used in making decisions about the significance of differences among sets of data, and statistical significance should be used as a criterion for deciding whether or not changes are needed in training programs or materials. However, statistically significant differences are not always practically significant; that is, the magnitude of the differences may not justify the time and expense required to change training programs or materials. For example, adding 10 hours of instruction to a training program may increase scores on the end-of-course test by a statistically significant amount, such as from a 70 percent to a 74 percent GO-rate. However, the trainers may not feel that this amount of improvement justifies the addition of 10 additional hours of instruction. They may feel that an improvement of at least 20 percentage points is needed to justify the additional instructional time. Techniques are available for selecting samples so that statistically significant differences are likely to have practical significance; a statistician should be consulted during the planning stage of data collection. Even in the absence of such techniques, decisions about the implications of feedback should be based upon practical as well as statistical considerations. Changes in training programs and materials should be made only when the results of analysis meet criteria for both statistical and practical significance.

In some cases, practical decision-making criteria may be somewhat arbitrary, at least during initial iterations of feedback collection. Criteria may need to be selected such that values falling outside of the acceptable range will indicate a problem serious enough to warrant modifications in training programs or materials. An example of such a criterion is the guideline which states that if more than 20 percent of the soldiers tested on a task receive first-time NO GO's, then the training program for that task is inadequate and must be revised. A criterion such as this can be entered into a computerized data management system so that values falling outside the acceptable range are automatically listed, thus providing an easy way of identifying potential problem areas. Many of the criteria used in decision-making based on feedback are admittedly arbitrary, but to some extent so are those used in statistical analysis. The best available decision-making approach is to use criteria which are determined as objectively as possible, consistently applied, and modified based upon experience and feedback.

One reason for identifying feedback by the time frame during which it was collected is to allow changes occurring over time to be detected. The identification of trends in data is a criterion that is often used as the basis for making changes in training programs and materials. But interviews of users of feedback indicate that the identification of trends is frequently based on general feelings that recent data represent a significant departure from data obtained in the past. If training is going to be modified on the basis of trends, then the minimum requirement is to collect feedback on the area of concern on three or more separate occasions and plot the results graphically as evidence that a trend exists. Examination of the data using decision-making criteria such as those discussed above should indicate whether a sufficient trend exists to require a reaction. Statistical procedures are available for testing the nature and significance of trends. However, the complexity of these procedures and the assumptions implicit in them limit their use in the analysis of feedback. If such procedures are used for identifying trends, a knowledgeable statistician should be consulted. It should also be noted that trends other than those occurring over time may be of interest during analysis of feedback. For example, a trend toward increasing gunnery scores with increasing hours of practice on a gunnery training simulator and similar relationships will be useful information for training developers. The points made above for temporal trends apply to all types of trends of interest in feedback analysis.

Reporting of Data

The processing of feedback should not stop with the analysis and interpretation of results. In order to insure that needed changes are made in training programs and materials, data summaries and resulting conclusions and recommendations must be presented to appropriate decision-makers. Reports of the results of feedback collection have often in the past consisted of little more than a listing of the general problems discovered. To make a convincing argument for needed changes, these reports should present data to indicate the extent of the problems noted, and they should offer specific recommendations for problem solution. They should also provide information about how the data were collected, how many units and individuals were involved, and the time period

over which the data were gathered. This information should be presented in a clear concise manner, so that the changes needed and their probable impacts are readily apparent.

An automated data base capable of presenting information in a variety of formats can enhance the clarity, conciseness, and efficiency of the data presentation. The data management system should have the capability to quickly generate line graphs, histograms, and other simple pictorial representations of the data. It should also be able to generate tables, statistics, and summaries of major problems observed within a given time frame for inclusion in monthly or quarterly evaluation reports. These reports could be composed and stored on the computer, making them readily available to users through interfacing terminals. Users could also create their own reports for special purposes by selecting the appropriate data, analysis options, and presentation formats through the terminal. To support this capability, an appropriate command from a terminal should access a catalogue of the various types of data available in the data base and the options and procedures available for analyzing and displaying them. From this list the user could select the type of data and method of analysis to be employed. Subsequent instructions could be used to select the time frame and soldier population of interest. With such a system, the user could quickly retrieve the data needed in a form best suited for particular applications.

Follow-Up

The gathering and use of feedback is a continuous cyclic process that does not end when changes have been made in training programs or materials. Following such changes in response to feedback, it is important to measure their effect. This usually involves collection of the same sorts of feedback which led to original identification of the need for change. For example, if performance on the test given at the conclusion of a course in map-reading increases noticeably following the introduction of modifications to the instruction, then the indication is that the changes increased training effectiveness. The degree to which the improvement in performance can be attributed to the course modifications depends on keeping other factors constant, such as the skill and abilities of the soldiers being trained and the test content and administration procedures. Follow-up is necessary to insure that changes instituted on the basis of feedback are producing the intended outcomes. If negligible or unfavorable outcomes result, the feedback system itself may require modifications, since it is leading to ineffective changes.

As indicated in the first section of this handbook, DOES personnel frequently collect feedback and pass information along to training developers concerning problems with training programs and materials. Training developers are expected to take corrective actions, but they may not do so because they are not convinced that a change is necessary or they do not have enough information to know what specific changes are required. The result is that feedback is not used and has no impact. Unless the collectors of feedback actively follow up to determine what changes are made and what recommended changes are not, they may incorrectly assume that the feedback they provided was completely satisfactory

and resulted in appropriate actions. The ultimate solution to this problem lies in the achievement of continuous cooperation and communications between training developers and evaluators. Since such a situation does not presently exist, collectors of feedback should follow up information that they have provided to other agencies, in order to determine exactly what changes were made and their effects. A feedback management system must include a mechanism for monitoring itself to insure that it has the desired effects.

APPENDIX A

SAMPLE QUESTIONNAIRES

ARTEP TRAINER QUESTIONNAIRE

INSTRUCTIONS: The purpose of this questionnaire is to obtain information on your use of the Army Training and Evaluation Program for Mechanized Infantry/ Tank Task Force (ARTEP 71-2). The information you provide will help insure that ARTEP 71-2 truly serves the needs of field units. The questions below address the ARTEP document itself as well as the process by which you use this document. Provide answers based upon your unit's experiences in training and evaluation exercises planned and conducted by you or higher headquarters. Answer the questions for the level of unit to which you are assigned. That is, if you are a platoon leader or sergeant, answer for the platoon; if you are a company commander, XO, or 1SG, answer for the company; and if you are a battalion commander or S3, answer for the battalion.

First provide the background information requested by filling in the blanks below. Then answer each question by circling the letter corresponding to your response and writing brief responses, where appropriate. More than one response may be appropriate for some questions; these are indicated below. No attempt will be made to identify or evaluate you or your unit, so please answer each question as completely and honestly as possible.

Date: _____ Duty Position: _____ Rank: _____

Specialty Code/MOS: Primary _____ Secondary _____

Months in Duty Position: _____ Months in Present Unit: _____

Last Institutional Training Completed
(AOB, ANCOC, etc.):_____ Year Completed: 19___

1. What type(s) of armored vehicles does your unit have?

a. M48A5s	d. M551s	g. M106s
b. M60A1s	e. M1s	h. M901s
c. M60A3s	f. M113s	i. M3s

2. Where is your unit located?

a. Continental United States (CONUS)
b. Korea
c. Germany

3. What component does your unit belong to?

a. Active Army
b. National Guard
c. Reserve

4. Have you used ARTEP 71-2 or participated in an exercise derived from it during the past 6 months (circle both a and b, if appropriate)?

a. Yes, in my present unit.
b. Yes, in another unit.
c. No.

NOTE: If you circled c in question 4, do not complete the rest of this questionnaire, since many of the questions ask about your experiences with the ARTEP in the last 6 months.

5. In what form is ARTEP 71-2 available to you? (Circle as many responses as are appropriate.)

a. I have my own copy of the ARTEP.

b. I have copies of the pages of the ARTEP that are relevant to my unit.

c. I have cards or other papers that contain information from the ARTEP relevant to my unit.

d. The ARTEP is in a central location where I can use it. (Where?)

_____

e. I have been briefed on the ARTEP but do not have access to written ARTEP materials.

f. Other. (Please explain.) _____

_____

6. How often do you use the ARTEP materials that are available to you?

a. Daily
b. Weekly
c. Monthly
d. Once a year
e. Never

7. Which version of ARTEP 71-2 is available to you?

a. New version published in November 1981.

b. Draft version published in March 1981.

c. Earlier version.

8. Have you received training in how to use ARTEP 71-2? (Circle as many answers as are appropriate.)

a. Yes, in institutional training (AOB, AOAC, ANCOC, etc.).

b. Yes, in a Battalion Training Management System (BTMS) workshop.

c. Yes, from the chain of command in my present unit.

d. Yes, other. (Please explain.) _____

_____

e. No.

9. If you have received training in use of ARTEP 71-2, did it provide you with sufficient information so that you could use the ARTEP in training your unit?

a. Yes.

b. No. (Please explain.) _____

10. Is the terminology in ARTEP 71-2 clear to you?

a. Yes.

b. No. (Please explain.) _____

_____

11. Is the terminology in ARTEP 71-2 clear to your soldiers?

a. Yes.

b. No. (Please explain.) _____

_____

12. Are there any missions or tasks <u>not</u> included in ARTEP 71-2 which your unit must perform to be combat ready?

a. Yes. (Please list them.) _____

_____

b. No.

13. Are there any missions or tasks included in ARTEP 71-2 that should be deleted?

a. Yes. (Please list missions and why they should be deleted.) _____

_____

_____

b. No.

14. How many training and evaluation exercises from ARTEP 71-2 has your unit conducted or participated in during the past 6 months? _____

15. How many of these training and evaluation exercises were planned and evaluated by: (Circle as many answers as are appropriate and fill in blanks).

a. Platoon? _____

b. Company? _____

c. Battalion? _____

d. Brigade? _____

e. Division? _____

16. How many of your unit's training and evaluation exercises, during the past 6 months, actually used OPFOR? _____

If OPFOR was not used, why not? _____

_____

_____

17. How many of your unit's training and evaluation exercises, during the past 6 months, used MILES? _____

If MILES was not used, why not? _____

_____

_____

18. How many firing proficiency/live-fire exercises has your unit conducted during the past 6 months? _____

Which ones; in what mode (dry-fire, MILES, subcaliber, full-caliber)? _____

_____

_____

_____

_____

19. Which of the General Missions from ARTEP 71-2 has your unit integrated into training and evaluation exercises during the past 6 months? Circle as many as applicable.

a. Plan and control combat operations

b. Maintain operations security (OPSEC)

c. Perform tactical intelligence functions

d. Conduct NBC defense operations

e. Defend against air attack

f. Conduct sustaining operations

20. For your unit's training and evaluation exercises, during the past 6 months, who selected the missions and tasks (in terms of duty position)? _____

_____

_____

21. On the average, how much time have you had to plan each of your unit's training and evaluation exercises during the past 6 months? _____

If this has not proved sufficient, how much time do you need? _____

22. Do you use any documents in addition to ARTEP 71-2 in order to train to standard?

a. Yes. (Which ones?) _____

_____

_____

b. No.

23. Did the conditions listed in ARTEP 71-2 match those your unit actually experienced in training and evaluation exercises during the past 6 months?

a. Yes.

b. No. (Please explain.) _____

_____

_____

24. Was your unit trained and evaluated in accordance with the standards listed in ARTEP 71-2?

a. Yes.

b. No. (Please explain.) _____

_____

_____

_____

25. Were you required to repeat tasks and/or missions until your unit performed to ARTEP standards?

a. Yes.

b. No. (Please explain.) _____

_____

_____

_____

26. What type of feedback did you receive on your unit's mission performance?

a. None

b. General performance ratings

c. Detailed information on tasks performed well and poorly

d. Other. (Please explain.) _____

_____

_____

27. What type of feedback did you provide to your soldiers on their performance?

a. None

b. General performance ratings

c. Detailed information on tasks performed well and poorly

d. Other. (Please explain.) _____

_____

_____

28. Who provided feedback during training and evaluation exercises (in terms of duty position) for the following?

a. Platoon _____

b. Company _____

c. Battalion _____

d. Brigade _____

e. Division _____

29. How much time elapsed between mission performance and provision of feedback for exercises that were planned and evaluated at the following levels? (Place checkmarks (✓) in the appropriate boxes.)

	Immediate	2-24 Hrs	2-6 Days	7+ Days
a. Platoon				
b. Company				
c. Battalion				
d. Brigade				
e. Division				

30. For what purpose does your unit most often use ARTEP 71-2?

a. As a training tool

b. As a test

c. Other. (Please explain.) _____

_____

31. List any other comments that you have about ARTEP 71-2: _____

_____

_____

_____

_____

_____

_____

_____

_____

PROFICIENCY ESTIMATION QUESTIONNAIRE - 19K TANK COMMANDER

Today your performance of the tasks listed below will be tested. In order to get a GO on any task, you will have to perform all the steps on it properly. The equipment and manuals you need to perform each task will be available. Before you are tested, we want to see if you can accurately predict how well you will do on each task. Write in your name and the other information requested below, and then predict how you will do on the test by circling a letter beside each task indicating your response. The letters stand for the following responses:

Y -- Yes, I can do this task and will get a GO on it.
? -- I don't know whether I can do this task or not.
N -- No, I can't do this task and will get a NO GO on it.

Name* _____ Grade _____ Unit _____

How many months have you been an M1 TC? _____

		Circle one:		
_____	Perform before-operations checks and services on an M1 tank.	Y	?	N
_____	Perform gunner's and loader's prepare-to-fire checks and services on an M1 tank.	Y	?	N
_____	Perform tank commander's prepare-to-fire checks and services on an M1 tank.	Y	?	N
_____	Boresight the main gun on an M1 tank within 10 minutes.	Y	?	N
_____	Boresight a caliber .50 M2 HB machinegun within 20 minutes.	Y	?	N

Now we want you to do one more thing on this page. Rank how well you can do the five tasks above by putting the numbers 1 through 5 in the blanks to the left of each task. Put one number in each blank and use each number once. Put 1 beside the task that you can do the best, 2 beside the task that you can do next best, and so on to 5 beside the task that you expect to do the worst on. **THEN** GO ON TO THE NEXT PAGE.

*NOTE: Names will be used only to match predictions with test results and they will be removed once the data are tabulated. The information you provide will not be used to evaluate you or your crew members.

TC'S ESTIMATIONS OF GUNNER'S PROFICIENCY

Now that you've predicted how you're going to do on the test today, we want you to predict how the members of your crew will do. On this page you are to predict how your gunner will do. The tasks he will be tested on are listed below. Write in your gunner's name and the other information requested below, and then predict how he will do on the test by circling a letter beside each task indicating your response. The letters stand for the following responses:

Y -- Yes, my gunner can do this task and will get a GO on it.
? -- I don't know whether my gunner can do this task or not.
N -- No, my gunner can't do this task and will get a NO GO on it.

Gunner's Name_____

How many months has he been your gunner? _____

		Circle one:		
_____	Perform before-operations checks and services on an M1 tank.	Y	?	N
_____	Perform gunner's and loader's prepare-to- fire checks and services on an M1 tank.	Y	?	N
_____	Perform tank commander's prepare-to-fire checks and services on an M1 tank.	Y	?	N
_____	Boresight the main gun on an M1 tank within 10 minutes.	Y	?	N
_____	Boresight a caliber .50 M2 HB machinegun within 20 minutes.	Y	?	N

Now rank how well your gunner can do the five tasks above by putting the numbers 1 through 5 in the blanks to the left of each task. Put a 1 beside the task he can do best, a 2 beside the task he can do next best, and so on like you did for yourself. THEN GO ON TO THE NEXT PAGE.

TC'S ESTIMATIONS OF DRIVER'S PROFICIENCY

Now you are to predict how well your driver will do on today's test, just like you did for your gunner on the previous page. The tasks he will be tested on are listed below. Write in your driver's name and the other information requested below, and then circle your response for each task.

Y -- Yes, my driver can do this task and will get a GO on it.
? -- I don't know whether my driver can do this task or not.
N -- No, my driver can't do this task and will get a NO GO on it.

Driver's Name_____

How many months has he been your driver? _____

	Circle one:
_____ Perform before-operations checks and services on an M1 tank.	Y ? N
_____ Prepare driver's station for operation on an M1 tank within 12 minutes.	Y ? N
_____ Start and stop the engine on an M1 tank within 10 minutes.	Y ? N
_____ Secure driver's station on an M1 tank within six minutes.	Y ? N
_____ Perform after-operations checks and services on an M1 tank.	Y ? N
_____ Set headspace and timing on a caliber .50 machinegun within 10 minutes.	Y ? N

Now rank how well your driver can do the six tasks above by putting the numbers 1 through 6 in the blanks to the left of each task. THEN GO ON TO THE NEXT PAGE.

TC'S ESTIMATIONS OF LOADER'S PROFICIENCY

Your final task is to predict how well your loader will do on today's test. The tasks he will be tested on are listed below. Write in your loader's name and the other information requested below, and then circle your response for each task.

Y -- Yes, my loader can do this task and will get a GO on it.
? -- I don't know whether my loader can do this task or not.
N -- No, my loader can't do this task and will get a NO GO on it.

Loader's Name_____

How many months has he been your loader? _____

	Circle one:		
_____ Perform before-operations checks and services on an M1 tank.	Y	?	N
_____ Perform operator maintenance on the 105mm breechblock assembly on an M1 tank within 14 minutes.	Y	?	N
_____ Clear an M240 coax machinegun to prevent accidental discharge on an M1 tank within 30 seconds.	Y	?	N
_____ Install/remove an M240 coax machinegun on an M1 tank within 10 minutes.	Y	?	N
_____ Perform operator maintenance on an M240 coax machinegun within 4 1/2 minutes.	Y	?	N
_____ Set headspace and timing on a caliber .50 machinegun within 10 minutes.	Y	?	N

Now rank how well your loader can do the six tasks above by putting the numbers 1 through 6 in the blanks to the left of each task.

APPENDIX B

SAMPLE INTERVIEW FORMS

INTERVIEW: TANK CREWMAN

Date _____

Position: GNR LDR DVR

Months in Position _____

Unit _____

1. How long have you been the (G, L, D) on this tank? _____

2. How long have you been an M1 (G, L, D)? _____

3. In the last two weeks, how many hours have you spent in:

   Scheduled Training _____ Unscheduled Training _____

   (Training refers only to those activities whose primary purpose is to teach MOS related skills and knowledges.)

4. What tasks were trained?

5. Were tasks trained to standard? Yes____ No____

   If yes, what were the standards?

6. Who did the training? TC Plt Sgt Plt Ldr Other_____

7. Did you actually get to perform each task? Yes____ No____

8. Were you told how well you did on each task? Yes____ No____

9. If you had trouble, were you given additional training on that task? Yes____ No____

10. What other activities took up time?

    _____ Maintenance

    _____ Inspections

    _____ Details

    _____ Other (specify)

INTERVIEW: TANK CREWMAN

Date _____

Unit _____

11. Have you used the battalion learning center in the last two typical garrison weeks? Yes____ No____

If yes, what lessons? If no, why not?

During duty hours _____ After duty hours _____

As a member of your crew _____ On your own _____

Was there anyone there to help you? Yes____ No____

12. Have you used the MOS library in the last two typical garrison weeks? If not, why not? Yes____ No____

13. Are you a graduate of: BAT _____
(Which MOS, date)

NET CONUS _____ NET USAREUR _____
(Duty position in which trained, date)

14. What was your first duty position following graduation from BAT?

GNR_____ LDR_____ DVR_____ Other_____

15. Are you currently enrolled in ACCP? Yes____ No____

If yes, what course?

What other ACCP courses have you taken?

16. (If on tank, ask to see -10)

Present?_____ Used Appearance?_____ Up to Date_____
(Indicate Vol 1, 2, 3)

(If not on tank) Do you have a -10 on the tank? Yes____ No____

In the last two weeks, for which tasks was the -10 used?

INTERVIEW: TANK CREWMAN

Date _____

Unit _____

17. Do you have your own copy of the SM? Yes____ No____

What tasks have you looked at in the last two weeks?

18. What problems have you had in driving, loading, firing, your tank?

19. What problems have you had in maintaining your tank?

20. What tasks should you have been trained on in BAT/NET that you weren't?

Have you received training on these in the unit? Yes____ No____

If yes, which ones?

21. Is there anything you would like to add?

Interviewer _____

INTERVIEW: ARMOR COMPANY CDR, XO, and 1SG

Date_____

Position: CDR XO 1SG Unit_____

1. How long have you been in your present duty position in this company? _____

2. a. Are you a graduate of: AOB _____ (date) AOAC _____ (SC, date)

ANCOC _____ (MOS, date) MASTER GNR _____ (MOS, date)

NET CONUS _____ NET USAREUR _____ (duty position in which trained, date)

b. What other formal Army schooling have you had?

3. In the last two weeks, how many hours has your company spent in:

Scheduled training _____ Unscheduled training _____

Individual training _____ Collective training _____

(Training refers only to activities whose primary purpose is to teach MOS related skills and knowledges.)

4. What activities, other than training, has your company been involved in during the last two typical garrison weeks?

5. During the last two garrison weeks, how many hours did you personally spend supervising or monitoring training? _____

INTERVIEW: ARMOR COMPANY CDR, XO, and 1SG

Date_____

Unit_____

6. What individual tasks were trained during the last two weeks?

7. Were tasks trained to standard? Yes____ No____

   If yes, where did the standards come from?

8. Who conducted most of this training? _____ (duty position)

9. How were these tasks selected for training?

10. During the last quarter, what collective training has your company participated in? Where was this training conducted?

11. a. What problems came up during this training?

    b. What was done to address these problems?

INTERVIEW: ARMOR COMPANY CDR, XO, and 1SG

Date_____

Unit_____

12. Do personnel in your company have:

Trainer's Guide	Yes_____	No_____
Crew Drills	Yes_____	No_____
Soldier's Manuals	Yes_____	No_____
Job Books	Yes_____	No_____

13. Does your company use written training plans? Yes_____ No_____

If yes, in what form (lesson plans, T&EO's, etc.)? _____

_____

(ask to see one)

14. What other materials do you use to manage, conduct, or evaluate training in your company?

15. What training support materials do you need that you don't have?

16. a. What are the hardest tasks for crewmen in your company to perform?

b. What makes these tasks hard?

17. Do you receive satisfactory support from organizational maintenance? Yes_____ No_____

If no, explain.

INTERVIEW: ARMOR COMPANY CDR, XO, and 1SG

Date_____

Unit_____

18. a. What percentage of your soldiers have transferred into the company during the last quarter? _____

    b. What percentage of your soldiers have changed duty assignments within the company during the last quarter? _____

19. What are the operational constraints on training in your company?

20. a. What are your training goals for your company?

    b. What problems have you had in meeting your goals?

    c. What can the Center/School do to help you meet your goals?

21. a. Do you participate in battalion-level planning and decision-making? Yes_____ No_____

    If yes, in what way?

INTERVIEW: ARMOR COMPANY CDR, XO, and 1SG

Date_____

Unit_____

b. How do you involve subordinates in company-level planning and decision-making?

22. If you could change anything you wanted, what would you change to make training better?

23. Is there anything else you would like to add?

Interviewer_____

APPENDIX C

SAMPLE OBSERVATION WORKSHEETS

TRAINING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Training_____

Trainer(s)_____ Trainee(s)_____

Training Site_____ Training Started (Time)_____

OBSERVATIONS	YES	NO	NA	COMMENTS
1. Were the soldiers told the training objectives to include tasks, conditions and standards?				
2. Did the trainer define new terms and identify new equipment for the soldiers?				
3. Did the trainer demonstrate how the tasks and subtasks are performed?				
4. Was the demonstration conducted in such a way that soldiers could see, hear and comprehend what the trainer was doing?				
5. Did all soldiers practice each task and subtask?				
6. During practice, did all soldiers perform the tasks to standard unassisted?				
7. What steps gave soldiers the most trouble during practice?				
8. Did every soldier practice under the direct supervision of the trainer(s)?				
9. Were soldiers told what they were doing right or wrong during practice?				
10. What training equipment was used?				
11. Was there enough training equipment for the number of soldiers being trained?				
12. Did the training equipment work properly?				

TRAINING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Training_____

OBSERVATIONS	YES	NO	NA	COMMENTS
13. Were training aids or materials used during this training? List those used.				
14. Were job aids handed out during this training? Which ones?				
15. Did the training include use of the job aids?				
16. Did the training progress from explanation to demonstration to practice?				
17. Did the trainer answer the soldiers' questions to their satisfaction?				
18. Was the training site adequate from the standpoint of range facilities, space, support equipment and noise?				
19. Was the training site arranged so that the trainer could be seen and heard by all including yourself?				
20. Was the weather so bad as to distract from training?				
21. For scheduled training, was the training time somewhat shorter or longer than the time allotted?				
22. Were there stated or written training objectives for this training?				
23. Was there a written or stated standard of performance?				
24. Did the training follow a training plan? Which one?				
25. Was a test held?				

Training Ended (Time)_____

Observer_____

TRAINING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Training_____

Trainer(s)_____ Trainee(s)_____

Training Site_____ Training Started (Time)_____

	YES	NO	NA	COMMENTS
Training Objectives				
1. Was the lesson based on stated or written objectives?	Y	N	NA	
2. Did they include ...				
- tasks?	Y	N	NA	
- conditions?	Y	N	NA	
- standards?	Y	N	NA	
Lesson Presentation				
3. Did the trainer ...				
- review training objectives?	Y	N	NA	
- define new terms?	Y	N	NA	
- identify new equipment?	Y	N	NA	
- answer questions to students' satisfaction?	Y	N	NA	
- demonstrate tasks?	Y	N	NA	
Demonstration				
4. Could the soldiers ...				
- see the demo?	Y	N	NA	
- hear the demo?	Y	N	NA	
- follow the demo?	Y	N	NA	

TRAINING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Training_____

	YES	NO	NA	COMMENTS
Practice				
5. Did all soldiers ...				
- practice subtasks separately?	Y	N	NA	
- combine subtasks into whole-task practice?	Y	N	NA	
- perform tasks to standard unassisted?	Y	N	NA	
6. Did an instructor ...				
- directly supervise each soldier's practice?	Y	N	NA	
- tell soldiers what they were doing right and wrong?	Y	N	NA	
7. Did certain steps give soldiers more trouble? Which steps?	Y	N	NA	
Training Equipment and Materials				
8. List any of the following used during training ...				
- training equipment?	Y	N	NA	
- training aids?	Y	N	NA	
- job aids?	Y	N	NA	
9. Did the training equipment work properly?	Y	N	NA	

TRAINING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Training_____

	YES	NO	NA	COMMENTS
10. Did the training ...				
- focus on the job aids?	Y	N	NA	
- proceed from explanation to demonstration to practice?	Y	N	NA	
- follow a training plan? Which one?	Y	N	NA	
Training Environment				
11. Was the training site adequate with regard to ...				
- range facilities?	Y	N	NA	
- support equipment?	Y	N	NA	
- space?	Y	N	NA	
- noise distractions?	Y	N	NA	
12. Did discomfort due to poor weather distract from training?	Y	N	NA	

Training Ended (Time)_____

Observer_____

81

ORGANIZATIONAL MAINTENANCE OBSERVATION WORKSHEET Date_____

Maintenance Site_____ Unit_____

Time of Visit_____ Visit Announced_____ Visit Unannounced_____

OBSERVATIONS	YES	NO	NA	COMMENTS
1. Were any mechanics present in the motor pool? How many?				
2. Were any mechanics engaged in maintenance activities? How many?				
3. Did you observe mechanics engaged in any other activities? What activities?				
4. Was any maintenance being performed on the M1 tanks in the motor pool? What was happening?				
5. Was there marked damage to the external bodies of the M1 tanks?				
6. Did you see copies of the -20 on the premises?				
7. Did the -20's show signs of use (e.g., greasy fingerprints)?				
8. Were the tools in the tool boxes clean?				
9. Were tools being borrowed or shared among maintenance stations?				
10. Was the STE/M1 in use?				
11. Was there a separate tool room?				
12. Was there a cleaning area for cleaning tools and parts with cleaning rags, solvents, etc.?				
13. Was maintenance supervision available at the maintenance site?				

ORGANIZATIONAL MAINTENANCE OBSERVATION WORKSHEET Date_____

Maintenance Site_____ Unit_____

OBSERVATIONS	YES	NO	NA	COMMENTS
14. Were supervisors actively involved in the maintenance activities?				
15. Was any training being conducted during your visit?				
16. Were work flow charts posted?				
17. Were there recent entries on the work flow charts?				
18. Were any crew members observed ordering parts from a 20 P manual?				

Observation Ended (Time)_____

Observer_____

TESTING OBSERVATION WORKSHEET Date_____ Unit_____

Name or Type of Examination_____

Examiner(s)_____ Examinee(s)_____

Testing Site_____ Test_____ Retest_____

Testing Started (Time)_____

OBSERVATIONS	YES	NO	NA	COMMENTS
1. Did the examiner read the test instructions to the soldiers?				
2. Were the test instructions stated clearly enough for the soldiers to know what they were expected to do?				
3. Were pass/fail standards clearly explained such that soldiers would know when they performed correctly?				
4. Were all tasks that were trained also tested?				
5. Were soldiers tested on any tasks that were not taught?				
6. Were the testing conditions the same as the training conditions?				
7. Were the standards used during training also used to score test performance?				
8. Did the examiner help or prompt the soldiers during the test?				

Testing Ended (Time)_____

Observer_____

APPENDIX D

REFERENCES

General

Bakken, D. & Bernstein, A. L. A systematic approach to evaluation. Training and Development Journal, 1982, 36 (8), 44-51.

Burnside, B. L. Field performance feedback - A problem review (Research Report 1323). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Aug 1981.

Burnside, B. L. Subjective appraisal as a feedback tool (Draft Technical Report). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, May 1982.

Hall, E. R., Lam, K., & Bellomy, S. G. Training effectiveness assessment: Current military training evaluation programs (TAEG Report No. 39, Volume I). Orlando, FL: Training Analysis and Evaluation Group, Dec 1976. (AD A036 517)

Hall, E. R., Rankin, W. C., & Aagard, J. A. Training effectiveness assessment: Problems, concepts, and evaluation alternatives (TAEG Report No. 39, Volume II). Orlando, FL: Training Analysis and Evaluation Group, Dec 1976. (AD A036 518)

Kristiansen, D. M. A job aid for modifying ineffective or inefficient training programs (Research Product 81-17). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Sep 1981.

Kristiansen, D. M. & Witmer, B. G. A job aid for the sy natic evaluation of lesson plans (Research Product 81-15). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Jan 1982.

Lewy, A. (Ed.) Handbook of curriculum evaluation. New York: Longman, Inc., 1977.

Morris, L. L. & Fitz-Gibbon, C. T. How to measure program implementation. Beverly Hills, CA: Sage Publications, 1978.

Patton, M. Q. Qualitative evaluation methods. Beverly Hills, CA: Sage Publications, 1980.

Salinger, R. D. & Deming, B. S. Practical strategies for evaluating training. Training and Development Journal, 1982, 36 (8), 20-29.

Witmer, B. G. & Burnside, B. L. Feedback needs of training developers and evaluators (Draft Research Report). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Aug 1982.

Questionnaires

Dyer, R. F., Josephine, J. M., Wright, C. E., & Yudowitch, K. L. Questionnaire construction manual. Operations Research Associates, July 1976. (AD A037 815)

Hawes, J. A. 1979 Reserve force studies surveys: Description and evaluation of survey procedures. Santa Monica, CA: The Rand Corporation, Sep 1981. (AD A111 904)

Sudman, S. & Bradburn, N. M. Asking questions: A practical guide to questionnaire design. San Francisco, CA: Jossey-Bass, Inc., 1982.

US Army Missile and Munitions Center and School, Directorate of Evaluations and Standardization. Handbook for developing questionnaires for missile and munitions evaluation (MAME). Redstone Arsenal, AL: Author, Jan 1981.

US Army Test and Evaluation Command. Questionnaire and interview design (TECOM Pam 601-1, Vol 1). Aberdeen Proving Ground, MD: Author, July 1975. (Part 1). (AD A060 499)

Interviews

Gorden, R. L. Interviewing strategy, techniques, and tactics. Homewood, IL: Dorsey Press, 1969.

Hall, E. R. & Hughes, H., Jr. A structured interview methodology for collecting training feedback information (TAEG Report No. 92). Orlando, FL: Training Analysis and Evaluation Group, Dec 1980. (AD A097 671)

Patton, M. Q. Qualitative evaluation methods. Beverly Hills, CA: Sage Publications, 1980. (Chapter 7)

US Army Test and Evaluation Command. Questionnaire and interview design (TECOM Pam 601-1, Vol. 1). Aberdeen Proving Ground, MD: Author, July 1975. (Part 2) (AD A060 499)

Observations

Kristiansen, D. M. & Witmer, B. G. Guidelines for conducting a training program evaluation (Research Product 81-18). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Feb 1982.

Morris, L. L. & Fitz-Gibbon, C. T. How to measure program implementation. Beverly Hills, CA: Sage Publications, 1978. (Chapter 5)

Patton, M. Q. Qualitative evaluation methods. Beverly Hills, CA: Sage Publications, 1980. (Chapter 6)

Roid, G. H. & Haladyna, T. M. A technology for test-item writing. New York: Academic Press, 1982. (Chapter 5)

Witmer, B. G. A job aid for the structured observation of training (Research Product 81-16). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Sep 1981.

Testing

Denova, C. C. Test construction for training evaluation. New York: Van Nostrand Reinhold Company, 1979.

Osborn, W. C., Campbell, R. C., & Ford, J. P. Handbook for the development of skill qualification tests (Final Report P-77-5). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Nov 1977.

Pearlstein, R. B. & Swezey, R. W. Criterion-referenced measurement in the Army: Development of a research-based, practical, test construction manual (Technical Report 78-A31). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Sep 1978.

Roid, G. H. & Haladyna, T. M. A technology for test-item writing. New York: Academic Press, 1982.

Schulz, R. E., Hibbits, N., Wagner, H., & Seidel, R. J. On-line authoring aids for instructional design (Technical Report 387). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Jul 1979.

Swezey, R. W. & Pearlstein, R. B. Guidebook for developing criterion-referenced tests. Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, Aug 1975. (AD A014 987)